

**Untersuchungen der Populationsheterogenität
während der heterologen Proteinproduktion in
*Bacillus megaterium***

Von der Fakultät für Lebenswissenschaften
der Technischen Universität Carolo-Wilhelmina zu Braunschweig
zur Erlangung des Grades
einer Doktorin der Naturwissenschaften
(Dr. rer. nat.)
genehmigte
D i s s e r t a t i o n

von Karin Michaela Münch
aus Friedrichshafen

1. Referent: Professor Dr. Dieter Jahn
2. Referentin: Privatdozentin Dr. Simone Bergmann
eingereicht am: 10.08.2017
mündliche Prüfung (Disputation) am: 19.10.2017

Druckjahr 2018

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen

Muench, K., Müller, J., Wienecke, S., Bergmann, S., Heyber, s., Biedendieck, R., Muench, R. and Jahn D. (2015). Polar fixation of plasmids during recombinant protein production in *Bacillus megaterium* results in population heterogeneity. *Environ Appl Microbiol* 81, 5976–5986.

Tagungsbeiträge

Münch, K.¹, Wienecke, S. ¹, Münch, R., Biedendieck, R. and Jahn, D. (2013). Phenotypic heterogeneity during heterologous protein production in *Bacillus megaterium*. Progress report meeting 2013, Bad Staffelstein, Germany

Posterbeiträge

Münch, K.¹, Wienecke, S. ¹, Münch, R., Biedendieck, R. and Jahn, D. (2014). Phenotypic heterogeneity during heterologous protein production in *Bacillus megaterium*. SPP 1617 Progress report meeting 2014 / International Conference: Phenotypic heterogeneity and sociobiology of bacterial populations, Irsee, Germany

¹ geteilte Autorenschaft

» Wer mit Ungeheuern kämpft, mag zusehen, dass er nicht dabei zum Ungeheuer wird.

Und wenn du lange in einen Abgrund blickst, blickt der Abgrund auch in dich hinein. «

Friedrich Wilhelm Nietzsche

Für meine Familie
Lara, Werner und Richard Münch

Ein Dankeschön an alle, die in jeglicher Hinsicht eine echte Hilfe waren.

Inhaltsverzeichnis

Abkürzungsverzeichnis	1
1. Einleitung	4
1.1 <i>Bacillus megaterium</i>	4
1.2 Heterologe Proteinproduktion	5
1.3 Kulturheterogenität	9
1.4 Einzelzellanalyse	11
1.4.1 Zeitraffer-Fluoreszenzmikroskopie	11
1.4.2 Durchflusszytometrie und Zellsortierung	12
1.5 Zielsetzung	13
2. Material und Methoden	15
2.1 Geräte	15
2.2 Verbrauchsmaterial	16
2.3 Chemikalien und Reagenzien	17
2.4 Bakterienstämme und Plasmide	18
2.5 Mikrobiologische Methoden	19
2.5.1 Kulturmedium für <i>Escherichia coli</i>	19
2.5.2 Kulturmedium für <i>Bacillus megaterium</i>	20
2.5.3 Antibiotika	21
2.5.4 Allgemeine Lösungen	21
2.5.5 Lösungen für die Herstellung und Transformation von Protoplasten	22
2.5.6 Lösungen für FISH	24
2.6 Zellbiologische Methoden	24
2.6.1 Sterilisation	24
2.6.2 Kultivierung von Bakterien	24
2.6.3 Induktion der plasmidtragenden Stämme	25
2.6.4 Konservierung von Bakterien	26
2.7 Molekularbiologische Methoden	26
2.7.1 Zellzahlbestimmung und Ermitteln von Wachstumskurven	26
2.7.2 Extraktion von Nukleinsäuren	26
2.7.3 Herstellung chemisch kompetenter Zellen für die Transformation in <i>E. coli</i>	27
2.7.4 Herstellung chemisch kompetenter Zellen für die Transformation in <i>B. megaterium</i>	28
2.7.5 Transformation in <i>E. coli</i>	28
2.7.6 Transformation in <i>B. megaterium</i>	29
2.7.7 Plasmidkonstrukte	29

2.7.8 Detektion der Fluoreszenz	33
2.7.9 DNA-Microarray	33
2.7.10 Quantitative Real-Time PCR (qRT-PCR)	35
2.7.11 Fluoreszenzmikroskopie und konfokale Laser-Scanning-Mikroskopie	41
2.7.12 Durchflusszytometrie und FACS	43
2.7.13 FISH	44
2.8 Software	46
2.8.1 BioLector	46
2.8.2 Microarray	47
2.8.3 qRT-PCR	47
2.8.4 Time-Lapse Mikroskopie.....	49
2.8.5 Durchflusszytometrie.....	49
3. Ergebnis und Diskussion	50
3.1 Makroskopische und mikroskopische Untersuchungen	50
3.2 FACS-Analysen von Subpopulationen, die viel und wenig eGFP produzieren	54
3.3 Analyse der Genexpression von <i>xylR</i> , <i>xylT</i> und <i>egfp</i> in den Subpopulationen mittels qRT-PCR	54
3.3.1 Reinheitsbestimmung der RNA	55
3.3.2 Ermitteln der Referenzgene	55
3.3.3 Relative Quantifizierung	59
3.3.4 Plasmid Copy Number (PCN)	60
3.4 Zeitraffer-Fluoreszenzmikroskopie	62
3.4.1 Entstehung von stark- und schwachproduzierenden Zelllinien	62
3.4.2 Einfluss des Xylose-Repressors auf die eGFP-Produktion	63
3.4.3 Zweikanal-Fluoreszenzmikroskopie zur Visualisierung von eGFP und XylR-mCherry	65
3.4.4 Zelllinienanalyse der Verteilung von eGFP und XylR-mCherry-Plasmid- komplexen	67
3.5 Plasmidlokalisierung mittels Fluoreszenz <i>in situ</i> Hybridisierung (FISH)	75
3.6 Durchflusszytometrie	78
3.6.1 Dynamik der Populationsheterogenität während der heterologen Protein- produktion von eGFP	79
3.6.2 Messung der Vitalität während der heterologen Proteinproduktion von eGFP in A5- und LB-Medium	88
3.6.3 Einfluss der Plasmidverteilung auf die Proteinproduktionsdynamik	99
3.7 Microarray	104
3.8 Wachstumskurven und mikroskopische Populationsheterogenität im Vergleich	113
3.8.1 Wachstum des Wildtyps im Vergleich zu den plasmidtragenden Stämmen ...	113
3.8.2 Mikroskopische Fluoreszenzaufnahmen diverser plasmidtragender Stämme	115

4. Zusammenfassung	117
4.1 Bistabilität wird offensichtlich nicht durch die positive Rückkopplung über die Xylose-Transporter vermittelt	117
4.2 Populationsheterogenität wird durch die ungleichmäßige Plasmidverteilung belegt	118
5. Summary	121
6. Ausblick	122
7. Referenzen	123
8. Anhang	134
8.1 Filme	134
8.2 R-Skripte zur Datenauswertung	135
8.2.1 qRT-PCR	135
8.2.1.1 R-Skript zur Berechnung der Effizienz und des Cq-Wertes, sowie zum Erstellen der Boxplots	135
8.2.1.2 R-Skript zum Ermitteln der Referenzgene	141
8.2.2 Microarray	142
8.2.2.1 Auswertung des DNA-Microarrays Teil 1	142
8.2.2.2 Auswertung des DNA-Microarrays Teil 2	146
8.2.3 Auswertung der differentiell exprimierten Gene	149
8.3 Auszug der differentiell exprimierten Gene	150
Danksagung	166

Abkürzungsverzeichnis

A_{λ}	Absorption bei der Wellenlänge λ in nm
AB3	Antibiotic Medium No. 3
Abb.	Abbildung
Amp	Ampicillin
AP	Alkaline Phosphatase
AU	Arbitrary Units (willkürliche Einheiten)
bp	Basenpaare
BSA	Bovine Serum Albumin
ca.	circa
cAMP	cyclisches Adenosinmonophosphat
CAP	Catabolite Activator Protein
Carb	Carbenicillin
CcpA	Catabolite Control Protein A
cDNA	komplementäre Desoxyribonucleinsäure
CIP	Calf Intestinal Alkaline Phosphatase
CP	Crossing Point
C_q	quantification Cycle
<i>cre</i>	Catabolite Response Element
Ct	Cycle Threshold
Cy	Cyanine
DEPC	Diethylpyrocarbonat
DNA	Desoxyribonucleinsäure
dNTP	Desoxyribonukleosidtriphosphat
dsDNA	double stranded DNA
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
DTT	Dithiotreitol
EDTA	Ethylendiamintetraacetat
eGFP	enhanced Green Fluorescent Protein
EMK	Erlenmeyerkolben
evtl.	eventuell
FACS	Fluorescence Activated Cell Sorting
FDM	First Derivative Maximum

for	forward
FSC	Frontal Scatter
g	Gramm (Gewicht) oder Erdbeschleunigung (Zentrifugation)
GFP	Green Fluorescent Protein
GOI	Gene of Interest
HEPES	2-(4-(2-Hydroxyethyl)-1-pieraziny)-ethansulfonsäure
HKG	Housekeeping Genes
H ₂ O _{deion}	deionisiertes Wasser
kb	Kilobasenpaare
lac	Lactose
LB	Lysogeny Broth
m	milli
M	molare Masse
M	M-Wert, Stabilitätsfaktor in geNorm
μ	mikro
Mb	Megabasen
MCS	Multiple Cloning Site
min	Minuten
MIQE	Minimum Information for Publication of Quantitative Real-Time PCR Experiments
M _r	relative Molekülmasse
n	nano
nm	nanometer
NRT	No Reverse Transcription Control
NTC	No Template Control
OD	optische Dichte
o/n	over night
ORF	Open Reading Frame
ori	Origin of Replication
PBS	Phosphate-buffered Saline
PCN	Plasmid Copy Number
PCR	Polymerase Chain Reaction
PEG 6000	Polyethylenglycol (zahlenmittlere molekulare Masse: 6000)
PI	Propidiumiodid

qRT-PCR	quantitative Real-Time PCR
RBS	Ribosome Binding Side
rep	Plasmide Replication Protein
rev	reverse
RIN	RNA Integrity Number
RNA	Ribonukleinsäure
RNase	Ribonuklease
ROX	6-Carboxy-X-Rhodamine
rpm	Revolutions per Minute
RT	Raumtemperatur
S	Svedberg
SDM	Second Derivative Maximum
sek	Sekunde
SSC	Side Scatter
ssDNA	single stranded DNA
Tab.	Tabelle
TAE	Tris-acetate/EDTA
TE	Tris-EDTA
TEMED	Tetramethylen diamine
Tet	Tetracyclin
TOF	Take-Off Point
TRIS	Tris(hydroxymethyl)-aminomethane
U	Unit
u. a.	unter anderem
ULS	Universal Linkage System
usw.	und so weiter
UV	Ultraviolett
V_{NF}	Variation des Normalisierungsfaktors in geNorm
v/v	Volume per Volume
W_R	Relative Werte in geNorm
w/v	Weight per Volume
z. B.	zum Beispiel

1 Einleitung

1.1 *Bacillus megaterium*

Mit einer enormen Größe von ca. $4 \times 1,5 \mu\text{m}$ (Bunk *et al.*, 2010) gehört das Gram-positive Bakterium *Bacillus megaterium* zu einem der größten bekannten Vertreter mit einem 100-fach größeren Volumen als *Escherichia coli* (Vary *et al.*, 2007). Der Größenvergleich zwischen beiden Organismen wird in Abbildung 1A deutlich. Erstmals wurde dieser Gigant 1884 von Anton De Bary beschrieben (De Bary, 1884). Die Größe seines Genoms beträgt ungefähr 5,1 Mbp und wurde von Eppinger *et al.*, 2011 sequenziert (Abb. 1B).

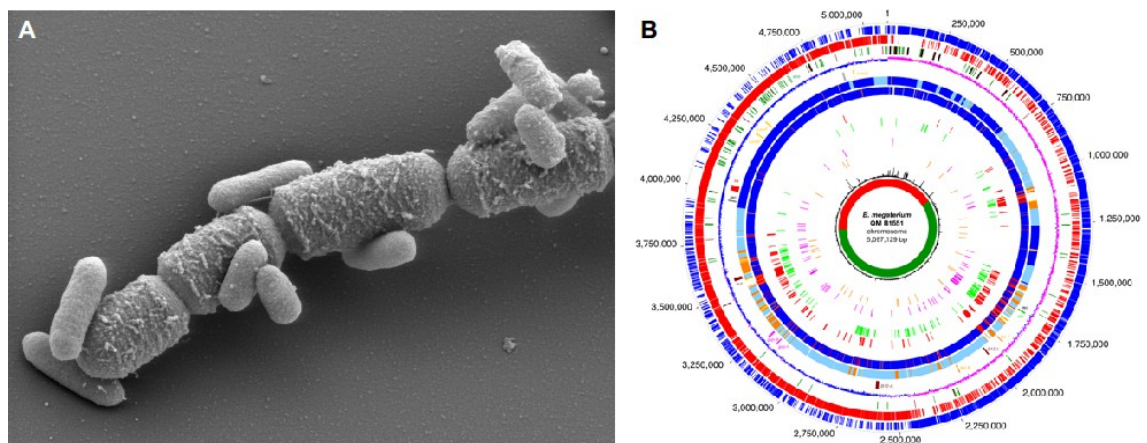


Abbildung 1: Morphologie und Genomsequenz von *Bacillus megaterium*. A) Die elektronenmikroskopische Aufnahme von M. Rohde (Helmholtz Zentrum für Infektionsforschung) verdeutlicht den Größenunterschied von *Bacillus megaterium* (große Kolonien) und *Escherichia coli* (kleine Zellen). B) Die Genomsequenz (Eppinger *et al.*, 2011) umfasst in etwa 5,1 Mbp.

Die *Bacillus*-Gruppe steht im Mittelpunkt vieler industrieller Prozesse, da sie im Bereich von mehreren Tonnen für die biotechnologische Herstellung von Amylasen, Antibiotika, Proteasen und Spezialchemikalien wie z. B. Pyruvat und Vitamin B12 eingesetzt wird (Vary *et al.*, 2007 und Biedendieck *et al.*, 2010b). *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus amyloliquefaciens* und *Bacillus megaterium* sind nur einige Vertreter für diese kommerziell wichtige Klasse von Bakterien (Liu *et al.*, 2013 sowie Dong and Zhang, 2014). Insbesondere *B. megaterium* ist aufgrund seiner

Apathogenität, mangelnden Endotoxine und alkalischen Proteasen sowie der leistungsstarken Proteinsekretion ein attraktiver Wirt für die heterologe Proteinproduktion (Biedendieck *et al.*, 2010a). Besonders im Bereich der Herstellung von Lebensmitteln, Kosmetikprodukten und Pharmaka besticht dieses Bakterium durch seine fehlenden Toxine. Wegen der nicht existierenden alkalischen Proteasen konnte während der heterologen Proteinproduktion so gut wie keine Degradation der Proteine beobachtet werden.

Ferner zeichnet sich *B. megaterium* durch eine hohe Plasmidstabilität aus (Meinhardt *et al.*, 1989 und Vary, 1994). Ein weiterer wichtiger Grund für die Nutzung dieses Bakteriums ist die Möglichkeit, auf verschiedenen Kohlenstoffquellen wachsen zu können, was die Kultivierung kostengünstig hält (Terpe *et al.*, 2006). Aufgrund seines enormen Spektrums an Substraten dient es zum Detoxifizieren des Bodens, da es selbst Stoffe wie Cyanide, Herbizide und persistente Pestizide metabolisiert. Bereits kleinste Spuren von Mangan reichen aus, um Sporen auszubilden. Das endosporenbildende Bodenbakterium bildet beim Wachstum meistens Ketten aus und besiedelt eine Vielzahl von Habitaten wie Seewasser, Reis, Milch, getrocknetes Fleisch oder Bienenhonig (Vary, 1994). Obendrein kann es auf einer Vielzahl von Oberflächen isoliert werden einschließlich klinischer Proben, Leder, Papier und Stein, aber auch im Kot der Kuh. *B. megaterium* wurde als Endophyt erkannt und in einigen Stämmen konnte sogar eine Stickstofffixierung nachgewiesen werden. Selbst in Abfällen der Fleischindustrie und der Petrochemie war es nachweisbar.

1.2 Heterologe Proteinproduktion

Zur Produktion von rekombinanten Proteinen wurde eine Reihe von Plasmiden konstruiert, deren Expressionssystem aus einem Xylose-induzierbaren Promotor P_{xylA} besteht, der seinerseits aus dem genomischen *xylABT*-Operon stammt (Ryguis and Hillen, 1991). Die Gene *xylA* und *xylB* codieren für Enzyme, die am Abbau von Xylose beteiligt sind, wohingegen *xylT* für einen Xylose-Transporter codiert. In Abwesenheit von Xylose wird die Expression des Operons durch den Xylose-Repressor *xylR* verhindert (Repression). In Gegenwart von Xylose bindet diese an den Repressor und

bewirkt damit eine Konformationsänderung, was die Repression aufhebt (Dereprimierung) und dadurch die Expression des Operons bewirkt. Eine schematische Darstellung dieses Systems zeigt Abbildung 2. Das *xylR*-Gen befindet sich stromaufwärts des *xylABT*-Operons, wird in divergenter Richtung transkribiert und negativ autoreguliert (Rygus *et al.*, 1991). Da *xylR* divergent zum Promotor P_{xylA} liegt und auch unter dessen Kontrolle steht, reprimiert XylR sich über negative Autoregulation selbst. Für die Konstruktion eines Xylose-induzierbaren Expressionssystems wurde das *xylR*-Gen mit seinen entsprechenden Promotoren P_{xylR} und P_{xylA} in ein frei replizierendes, natürlich vorkommendes Plasmid kloniert und durch die Einführung starker Promotorelemente optimiert (Malten *et al.*, 2006). Im Unterschied zum P_{xylA} -Promotor des Plasmids beinhaltet der genomische P_{xylA} -Promotor ein *cre*-Element als Bindestelle für den Katabolitrepresor CcpA. Dadurch wird verhindert, dass Xylose in Gegenwart von Fructose das *xylABT*-Operon induziert (Rygus and Hillen, 1992). Zwischen den verschiedenen Zuständen des Repressors XylR, der sowohl gebunden und frei vorliegen kann und demnach aktiv und inaktiv vertreten ist, liegen Gleichgewichte in der Zelle vor, die sich je nach der Anzahl der Plasmide durch die Gendosis verändern können. Das beschriebene System wurde als Shuttle-Vektor konstruiert und besitzt zur Replikation in *B. megaterium* das Gen *repU* und einen Replikationsursprung *oriU*. Für die Replikation in *E. coli* ist das Plasmid mit einem zweiten *ori* und dem Ampicillin-Resistenzgen *bla* ausgestattet. Zur Konstruktion dieses Vektors wurden ursprünglich die beiden Plasmide pBR327 aus *E. coli* und pBC16 aus *B. cereus* fusioniert, wodurch das Vorgängerplasmid pWH1508 entstand (Rygus *et al.*, 1991).

Das Tetracyclinresistenzgen ist als *tet(L)* klassifiziert und codiert für ein membranassoziiertes Efflux-System. Diese Genklasse wurde in den Plasmiden und Chromosomen verschiedener Gram-positiver Bakterien gefunden (Chopra and Roberts, 2001). Das letzte Gen des *xyl*-Operons *xylT* codiert für einen Protonen-Xylose-Symporter. Ein Terminator zwischen *xylA* und *xylB* entkoppelt seine Transkription vom P_{xylA} -Promotor und sorgt, wie bei Membranproteinen üblich, für eine niedrigere Transkriptionsrate (Schmiedel *et al.*, 1997).

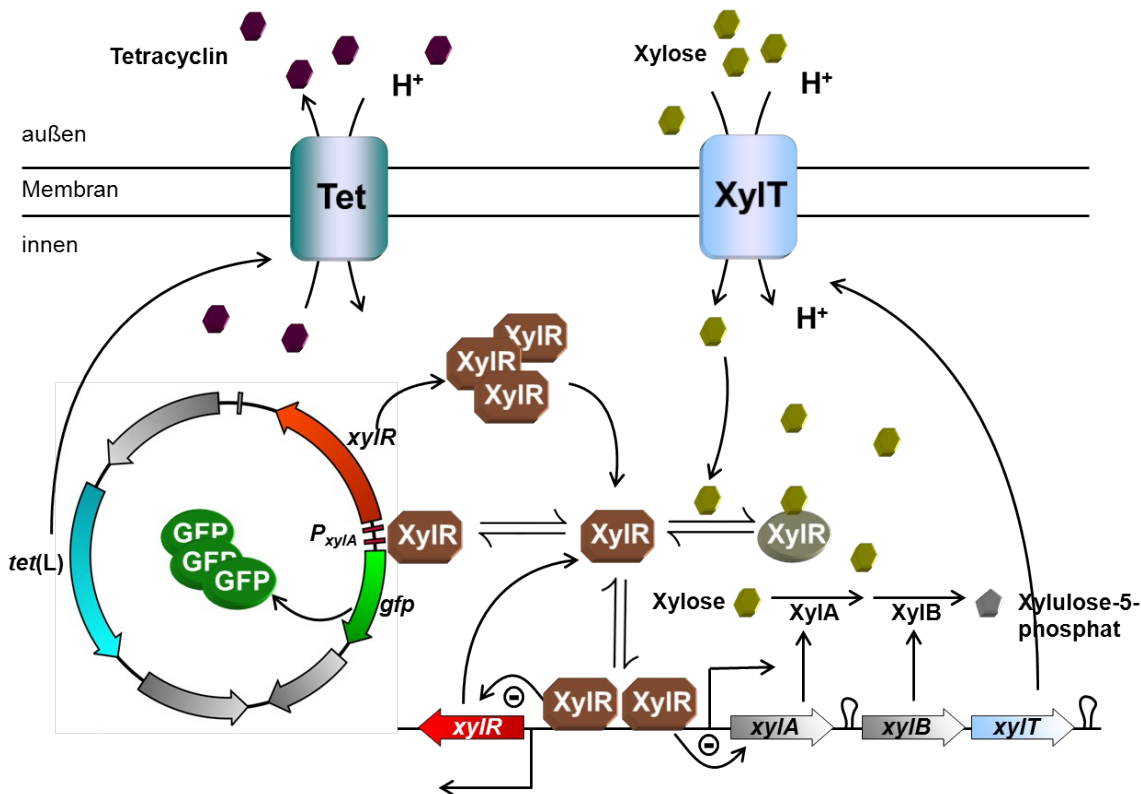


Abbildung 2: Schema des Xylose-induzierbaren Expressionssystems. Xylose gelangt über den entsprechenden Transporter XylT in die Zelle. In Gegenwart von Xylose bindet diese sowohl an den divergenten Xylose-Repressor (XylR) des Genoms, als auch des Plasmids und bewirkt dadurch eine Expression der dahinterliegenden Gene. Innerhalb des Genoms bedeutet dies die Expressierung des $xylABT$ -Operons einschließlich der autoregulierten Produktion von XylR. In Bezug auf den Vektor ist die Induktion sowohl mit einer erhöhte Menge an $xylR$, als auch mit einer immensen eGFP-Produktion verknüpft.

Ein Abkömmling dieses Shuttle-Vektors wird als pSSBm85 bezeichnet und enthält das mittels Fluoreszenzmessung leicht nachweisbare eGFP (enhanced Green Fluorescent Protein), das als Reportergen an den P_{xylA} -Promotor fusioniert wurde. Bereits bei normalem Licht erscheint eine eGFP-produzierende Kultur grünlich (Abb. 3A). Durch Anregung im Blaulichtbereich fluoresziert sie grün (Abb. 3B sowie Abb. 4).

Aufbauend auf pSSBm85 konnten weitere Vektorsysteme entwickelt werden, die bis zu 1,25 g/l eGFP produzierten (Stammen *et al.*, 2010a). Allerdings zeigte ein wesentlicher Anteil der Kultur auf Einzelzellebene eine Heterogenität bezüglich der Proteinproduktion auf (Abb. 4B).



Abbildung 3: Flüssigkulturen des Produktionsstammes pSSBm85. A) Bei Normallicht erscheint eine Zellsuspension, die bereits eGFP herstellt leicht grün (links im Bild), im Vergleich zu einer Suspension, die kein eGFP produziert (rechts im Bild). B) Wird eGFP im Blaulichtbereich angeregt, fluoresziert die induzierte Bakterienkultur grün.

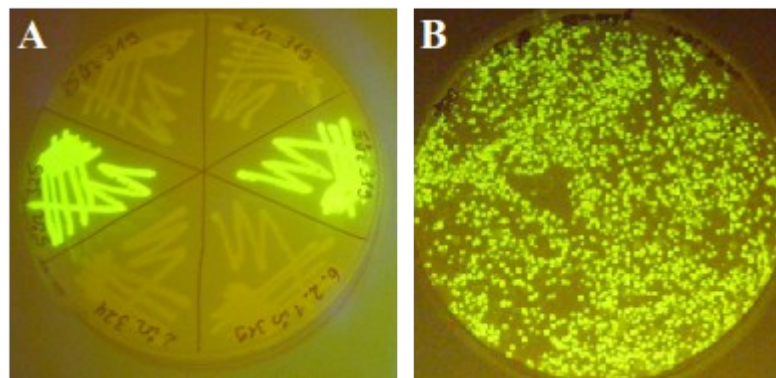


Abbildung 4: Stämme auf Festmedium. A) Kulturen, die bereits eGFP produzieren fluoreszieren grün (linke und rechte Seite). B) Wachsen sie in Form von Mikrokolonien, bilden sich nach der Induktion fluoreszierende und nicht fluoreszierende Subpopulationen aus.

Durchflusszytometrische Analysen einer genetisch identischen Kultur von *B. megaterium*, die in einem Bioreaktor gezüchtet wurde, dokumentieren selbst bei hochselektiven Bedingungen eine stabile Unterpopulation (ca. 30 %) von Niedrigproduzieren (Biedendieck *et al.*, 2007). Bei diesen Zellen handelte es sich um eine vitale Subpopulation. Somit stellt dieses Phänomen ein Beispiel für Kultur- oder Populationsheterogenität dar. Diese Resultate sind von großem kommerziellen Interesse, da das Wachstum dieser Wenigproduzierer die Gesamtproduktionsrate senkt und damit wertvolle Ressourcen während der heterologen Proteinproduktion verbraucht. Fluoreszenzmikroskopische Aufnahmen bestätigen die Kulturheterogenität durch das Vorhandensein verschiedener Subpopulationen (Abb. 5).

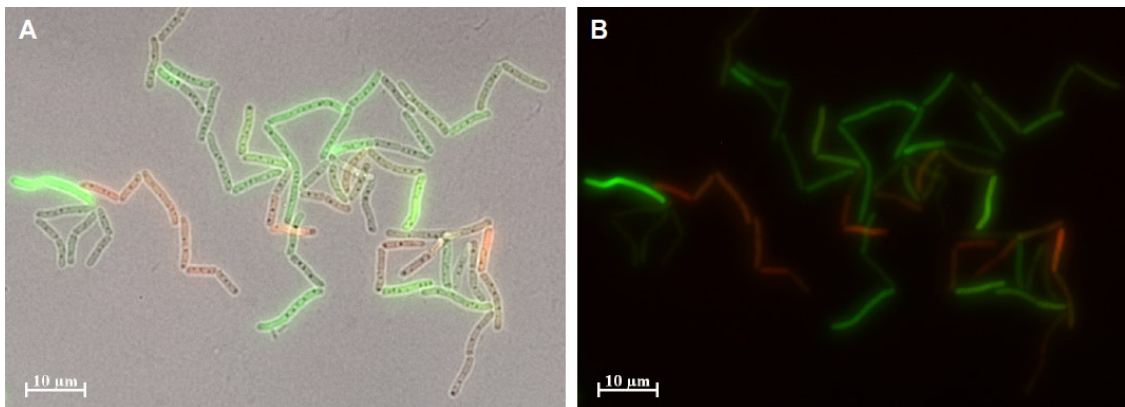


Abbildung 5: Fluoreszenzmikroskopische Aufnahmen zeigen die Kulturheterogenität einer Suspension von *B. megaterium*. Bei dieser Kultur handelt es sich um ein Derivat des Produktionsplasmids, das sowohl eGFP (grüne Fluoreszenz), als auch mCherry (rote Fluoreszenz) über ein XylR-mCherry-Fusionsprotein produziert. A) Zuerst sind die Kanäle für Hellfeld und den entsprechenden Fluoreszenzen dargestellt. B) Das nächste Bild zeigt den selben Ausschnitt ohne Hellfeldkanal.

1.3 Kulturheterogenität

Kulturheterogenität ist ein Phänomen, bei dem genetisch identische Zellen unterschiedliche Phänotypen ausprägen. Diese stabile Bildung zweier coexistierender Subpopulationen wird auch als Bistabilität bezeichnet. Bistabilität (lateinisch bi = zwei und lateinisch stabilis = standhaft, stabil) zeichnet sich durch die Möglichkeit aus, zwei stabile Zustände einnehmen zu können. In Zellen werden bistabile Zustände über genregulatorische Netzwerke mit positivem Feedback vermittelt (Dubnau and Losick, 2006). Dabei ist jeder Zustand stabil gegen äußere Signale. Sollten diese aber einen kritischen Wert übersteigen, kann es den Übergang in den anderen Zustand bewirken. Ausgelöst wird dies aufgrund von Umwelteinflüssen, molekularem Rauschen der Genexpression oder durch stochastische Variabilität (Kussell and Leibler, 2005 sowie Acar *et al.*, 2008). Die Bistabilität aufgrund stochastischer Verteilung wird auch als "Bet-hedging" Strategie bezeichnet und ist ein Beispiel für evolutionäre Anpassung an schwankende Umweltbedingungen. Sie stellt eine Art Lebensversicherungen für Bakterien dar (Ackermann *et al.*, 2008). Bei dieser Anpassungsstrategie befinden sich einzelne Zellen in physiologischen Zuständen, die nicht unbedingt vorteilhaft sind. Bei plötzlichen Umweltveränderungen können diese Organismen jedoch überleben und dadurch ein Fortbestehen der Population garantieren. Bistabile Systeme können

Hysterese (griechisch *hysteros* = hinterher, später) ausbilden, so dass ein Stadium auch nach dem Wegfall einer Stimulierung noch andauert. Die Entscheidung der Zelle welcher Zustand gewählt wird, hängt hierbei von der Vorgeschichte ab (Solopova *et al.*, 2013). Bekannte Phänomene für Bistabilität in Bakterien sind Sporulation, Persistenz bei der Antibiotikaresistenz und genetische Kompetenz. Ein besonderes Beispiel ist Leben und Tot auf der Zellebene innerhalb einer Zellkultur.

In Bakterien wurde Bistabilität erstmals vor über 50 Jahren bei der Expression des *lac*-Operons (Jacob and Monod, 1961), das für die Verwertung von Lactose verantwortlich ist, in *E. coli* gezeigt (Novick and Weiner, 1957 sowie Cohn and Horbita, 1959). Hier wird die Transkription des Operons durch den konstitutiv exprimierten *lac*-Repressor LacI unterdrückt. Eine Induktion findet durch Lactose statt und führt zur Derepression dieses Operons. Zusätzlich wird es über das Aktivatorprotein CAP aktiviert, dessen Aktivität wiederum von der Konzentration des cAMP, welches bei niedrigen Glucosekonzentrationen vorhanden ist, abhängt. Bei Glucosemangel reichern sich die Proteine des Phosphotransferasesystems (PTS) in phosphorylierter Form an und bewirken dadurch die Aktivierung der Adenylylcyclase, die die Bildung von cAMP aus ATP unter Abspaltung von Pyrophosphat katalysiert. Tritt Glucose in die Zelle ein, wird normalerweise der Phosphatrest von Phosphoenolpyruvat über das PTS auf die Glucose übertragen, wobei Glucose-6-phosphat (1. Schritt der Glycolyse) entsteht. Zusätzlich wird das Operon über die Katabolitrepression von einem Induktor (Lactose) kontrolliert. Liegt eines der Transportproteine (E IIA) des PTS unphosphoryliert vor, wird die Lactose-Permease gehemmt, so dass Lactose nicht in die Zelle transportiert werden kann und das *lac*-Operon somit inaktiv bleibt. Erst wenn keine Glucose mehr vorhanden ist, kann Lactose in die Zelle transportiert werden und an den *lac*-Repressor binden. Jedoch konnte als Reaktion auf den Induktor ein diskontinuierliches Umschalten der Expression des *lac*-Operons beobachtet werden. Man entdeckte Zellen, die große Mengen an β -Galactosidase produzierten und solche, bei denen nur wenig detektiert werden konnte.

Die beiden Modelle des Lactose- und Xylose-Operons weisen große strukturelle Ähnlichkeiten auf (Abb. 6). Zudem konnte die für das *lac*-Operon erläuterte Bistabilität auf eine positive Autoregulation des Lactose-Transporters zurückgeführt werden (Abb. 6A). Allerdings zeigte das *xyl*-Operon (Abb. 6B) eine negative Rückkopplung durch

Induktion des XylR-Repressors, der in diesem System freie Xylose wegtitriert und damit die in der Zelle vorkommenden Gleichgewichte verschieben kann.

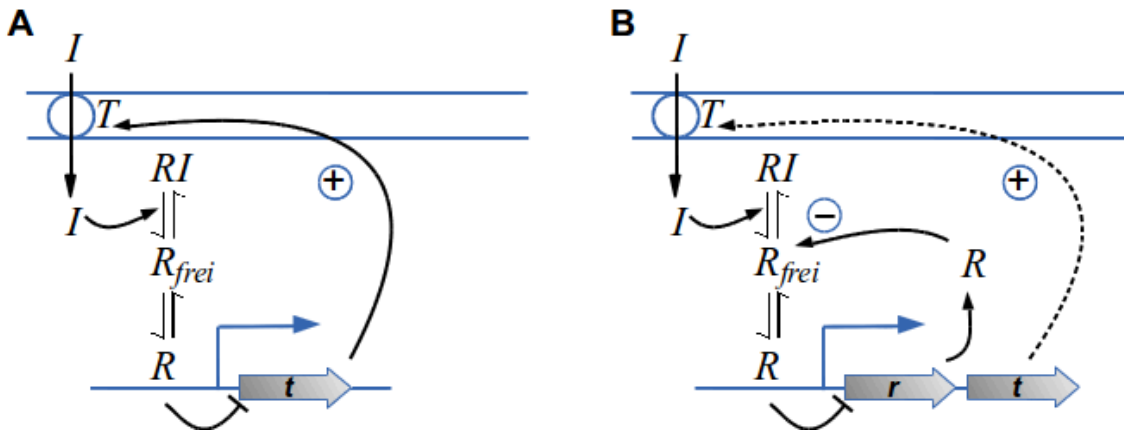


Abbildung 6: Positives Feedback der strukturell ähnlichen Operons des *lac*- und *xyl*-Systems. A) Im *lac*-Operon gelangt der Inducer (I) durch den Transporter (T) in die Zelle und bindet reversibel an den Repressor (R). Durch Bildung des Repressor-Inducer-Komplexes (RI) wird die Bindung zum Promotor aufgehoben (R_{frei}) und die Genexpression des Operons induziert. Beim *lac*-Operon wird der Transporter induziert, wodurch zunehmend mehr Inducer in die Zelle gelangt und eine positive Rückkopplung vermittelt wird. B) Auch beim *xyl*-Operon ist eine positive Rückkopplung über den Transporter XylT theoretisch denkbar (gestrichelte Linie). Allerdings wird dann zusätzlich mehr Repressor produziert, der den Inducer aus dem System wegtitriert und den Effekt der Bistabilität abschwächen könnte.

1.4 Einzelzellanalyse

Im Gegensatz zum traditionellen Ansatz, der eine gesamte Zellkultur betrachtet, beschreibt diese Analyse die Untersuchung einzelner Zellen. Um die phänotypische Heterogenität zu verstehen, ist es wichtig bedeutende Ereignisse auf Einzelzellebene zu analysieren, mit dem Ziel die molekularen Mechanismen aufzuklären.

1.4.1 Zeitraffer-Fluoreszenzmikroskopie

Bei der Zeitraffer-Fluoreszenzmikroskopie werden Bilderserien von sich teilenden Zellen angefertigt. Über Bildverarbeitungsprozesse können einzelne Zellen sowohl räumlich getrennt (Segmentierung), als auch zeitlich verfolgt (Tracking) werden.

Biologische Parameter wie Morphologie, Wachstum, Zellteilung, Lyse oder Migration können aus den Daten direkt abgeleitet und Zellstammbäume erstellt werden. Hierbei können Einzelzelleffekte über Fluoreszenzmarker visualisiert und quantifiziert werden. Der Vorteil dieser Methode ist die Erhaltung der zeitlichen Dynamik von Einzelzellprozessen und der Beziehung von Tochter- und Mutterzellen in einer wachsenden Population.

1.4.2 Durchflusszytometrie und Zellsortierung

Mit Hilfe von Durchflusszytometrie können einzelne Zellen in einem Flüssigkeitsstrom über ein optisches Messsystem analysiert werden. Die Zellsuspension wird hierbei durch eine Kapillare geleitet, so dass einzelne Zellen nacheinander von einem fokussierenden Laserstrahl beleuchtet werden. Das durch den Kontakt entstehende Streulicht oder Fluoreszenzsignal wird von einem Detektor ausgewertet. Das Vorwärtstreulicht (FSC = Forward Scatter) ist ein Maß für die Beugung des Lichts im flachen Winkel zur Auftreffrichtung, hängt vom Volumen der Zelle ab und gibt die ungefähre Zellgröße wieder. Das Seitwärtstreulicht (SSC = Side Scatter) ist ein Maß für die Brechung des Lichts im rechten Winkel und wird von der Granularität einer Zelle beeinflusst. Neben den Streulichtparametern können zusätzlich Fluoreszenzen detektiert werden (Abb. 7). Durch die Färbung mit DNA-interkalierenden Farbstoffen wie DAPI (4',6-Diamidin-2-phenylindol) oder PI (Propidiumiodid) kann die DNA-Menge bzw. die Vitalität einer Zelle gemessen werden.

Auch mit Fluorochromen gekoppelte Antikörper wie z. B. Alexa Fluor, Cy3 und Cy5 können zur Färbung verwendet werden. Das von den Zellen emittierte Licht ist dann proportional zur Menge ihrer gebundenen Antikörper. Während diese Methode bevorzugt für Oberflächen-Proteine geeignet ist, können intrazelluläre Proteine durch Fusion mit Fluoreszenzproteinen wie GFP oder mCherry detektiert werden.

Eine besondere Form der Durchflusszytometrie ist das Fluorescence-activated Cell Sorting (FACS). Die Zellsortierung wird dabei mittels Fluoreszenzmarkierung vorgenommen. Dabei wird die Größe des sich bildenden Tropfens so eingestellt, dass genau eine Zelle darin enthalten ist. Aufgrund der ausgestatteten Fluoreszenz wird dieser Tropfen zunächst elektrisch aufgeladen und anschließend im elektrischen Feld in das

entsprechende Gefäß sortiert.

Im Vergleich zur Zeitraffer-Fluoreszenzmikroskopie ist der Vorteil von Methoden die auf dem Prinzip der Durchflusszytometrie basieren darin begründet, dass eine große Anzahl von Zellen in kurzen Zeiträumen vermessen werden kann. Dies erlaubt einen statistisch signifikanten Überblick über die zu untersuchende Zellpopulation.

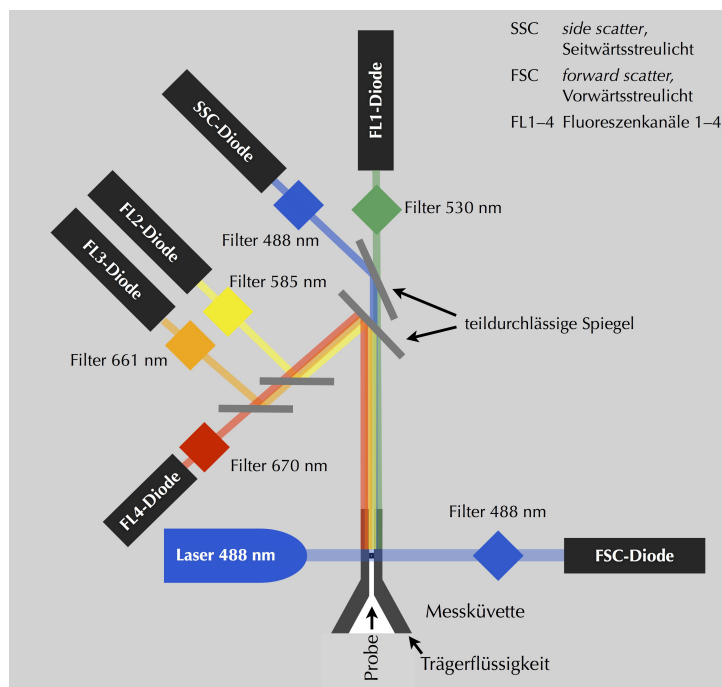


Abbildung 7: Prinzip eines Durchflusszytometers. Einzelne Zellen der Probe passieren mittels Kapillare einen Laserstrahl. Nachdem diese angestrahlt werden, entstehen sowohl Streulicht, als auch Fluoreszenzsignale. Entsprechend des gemessenen Winkels des Streulichts und der Fluoreszenzdetektion können Zellen durch Größe, Granularität und ihres Fluoreszenzsignals analysiert und optisch visualisiert werden. (Quelle: Wikipedia – Autor: FabianK)

1.5 Zielsetzung

In den letzten 10 Jahren wurde das Gram-positive Bakterium *Bacillus megaterium* systematisch für die Produktion von rekombinanten Proteinen im Gramm-pro-Liter-Maßstab weiterentwickelt. Als Expressionssystem wurde hierbei erfolgreich ein multicopy Plasmid mit Xylose-induzierbarem Zielgen eingesetzt. Weitere Studien während der Produktion von eGFP als Modellprotein haben gezeigt, dass trotz der

hohen Produktionseffizienz Subpopulationen mit nicht-produzierenden Zellen entstehen. In dieser Arbeit sollen diese Subpopulationen mit Techniken der Einzelzellanalyse analysiert, quantifiziert und die molekularen Grundlagen ihrer Entstehung aufgedeckt werden.

Zur Gewinnung weiterer Informationen über die molekularen Zusammenhänge der Zellheterogenität in der Gesamtpopulation sollte mittels einer Microarray-Analyse die differentielle Genexpression innerhalb der heterogenen Zellkultur untersucht werden. Dies sollte nach der Induktion mit Xylose in einer Zeitreihe durchgeführt werden.

Für die Analyse der heterogenen Population sollte die Zellkultur mittels Zellsortierung (FACS) separiert und die entsprechenden Komponenten des rekombinanten Genexpressionssystems mittels quantitativer Real-Time PCR charakterisiert werden. Wegen der strukturellen Ähnlichkeit zum *lac*-Operon sollte insbesondere der Einfluß des Xylose-Transporters auf die Kulturheterogenität kontrolliert und nachgewiesen werden, weil auch in diesem System eine positive Rückkopplung über den Xylose-Transporter theoretisch denkbar wäre. Durch Weiterkultivierung getrennter Subpopulationen sollte das Hysteresis- und Bistabilitätsverhalten analysiert werden.

Die Entstehung produzierender und nicht-produzierender Zelllinien in ihrer räumlichen und zeitlichen Dynamik sollte mit Hilfe von Zeitraffer-Fluoreszenzmikroskopie und Durchflusszytometrie beobachtet und quantifiziert werden. Zusätzlich sollten neue Vektoren konstruiert werden, um mit Hilfe eines XylR-mCherry-Fusionsproteins den Einfluss des Xylose-Repressors auf die Populationsheterogenität zu bestimmen.

Um stichfeste molekulare Informationen über den Einfluss der Plasmidkopienzahl zu gewinnen, sollte das Plasmid mit Hilfe eines XylR-mCherry-Fusionsproteins, ähnlich wie bei einem fluoreszierenden Repressor-Operator-System (FROS) zum *in vivo*-Tagging (Straight *et al.*, 1996) visualisiert werden und Aufschluss über die zelluläre Lokalisation und erste Einblicke in die Plasmidsegregation von multicopy Plasmiden geben. Die Verteilung und Lokalisation der Plasmide sollten zusätzlich über FISH-Experimente und DAPI Färbungen bestätigt werden.

2 Material und Methoden

2.1 Geräte

Alle Geräte, die für die Arbeit benötigt wurden, sind Tabelle 1 zu entnehmen.

Tabelle 1: Liste der verwendeten Geräte.

Gerät	Modell	Hersteller
Agarose Gel Dokumentation	GelDoc	Bio-Rad
Agarose Gel Elektrophorese	Agagel	Biometra
Autoclav	LVSY 50/70	Zirbus
Bioanalyzer	2100 Bioanalyzer	Agilent
Blaulicht-Transilluminator	Flu-O-Blu	Biozym
Digitalkamera	FinePix S700	Fujifilm
Durchflusszytometer	Cube 8	Sysmex Partec
Elektrophoresekammer	PerfectBlue	peqlab
FACS / Zellsortierer	FACSAria-II	Becton-Dickinson
	FACSAria-II SORP	Becton-Dickinson
Feinwaage	TE 2145	Sartorius
Fluoreszenzmikroskop	Axiovert 200 M	Zeiss
Gefrierschrank -20°C	-	privileg
Gefrierschrank -80°C	MDF U73V	Sanyo
Gradient Cycler	Tgradient	Biometra
Hybridisierungssofen	1013	Shel lab
Kühlschrank	-	privileg
Magnetrührer	MR 2002	Heidolph
Microarray-Scanner	high resolution scanner (G2565CA)	Agilent
Mikrobioreaktor	BioLector	m2p-labs
Mikroskop-Inkubator	Heating System 1 (ibiTC-025)	ibidi
Mikroskopkamera	Axiocam HR	Zeiss
Mikrowelle	-	LG
pH-Messung	pH-Meter CG 842	Schott

Photometer	BioMate3	Thermo Scientific
Pipetten	Research	Eppendorf
Power Supply	ENDURO	Labnet
		International
Real-Time PCR Detektionssystem	CFX96 Real Time System	Bio-Rad
	C1000 Thermal Cycler	
Reinstwasseranlage	MilliQ Synthesis A 10	Millipore
Schüttler	Aquatron	Infors HT
Sicherheitswerkbank	HERAsafe	Heraeus
Sonifikation	SONOPLUS	Bandelin
		electronics
Spectrofluorometer	FP-500	Jasco
Thermocycler	T Personal	Biometra
Thermomixer	Thermomixer compact	Eppendorf
UV-VIS Spektrophotometer	NanoDrop™ ND-1000	peqlab
Vortexer	Vortex Genie 2	ThermoScientific
Waage	572	Kern
Wasserbadschüttler	Aerotron AT15	Infors HT
Zeitschaltuhr	ZUEFTMINM	Düwi
Zentrifugen	Avanti J-E Centrifuge	Beckman Coulter
	Megafuge 1.0 R	Heraeus
	Mini Spin	Eppendorf

2.2 Verbrauchsmaterial

Allgemeine Verbrauchsmaterialien wie Küvetten, Pipettenspitzen, Reaktionsgefäße und Sterilfilter wurden von den Firmen BD Falcon, Brand, Eppendorf, GE Healthcare, Greiner bio-one, Invitrogen und Sarstedt bezogen. Spezielle Verbrauchsmaterialien sind in Tabelle 2 aufgeführt.

Tabelle 2: Spezielles Verbrauchsmaterial.

Material	Modell	Hersteller
Folien für Mikrotiterplatten	Gas-permeable sealing foil	m2p-labs
	Microseal 'B' Adhesive Seals	Bio-Rad
Objektträger	Poly-L-Lysine Slides	ibidi
Objektträgerschale	µ-Dish, 35 mm	ibidi
Mikrotiterplatten	FlowerPlate (48 well)	m2p-labs
	multiplate (96 well)	Bio-Rad
Reaktionsgefäße	Nennvolumina 1 ml	Sysmex Partec

2.3 Chemikalien und Reagenzien

Alle verwendeten Chemikalien wurden, sofern nicht anders angegeben, von den Firmen Carl Roth, Invitrogen, Merck Schuchard OHG, Serva oder Sigma-Aldrich bezogen.

Sämtliche Enzyme wurden von der Firma New England BioLabs hergestellt.

Diverse Kits und spezielle Reagenzien mit denen gearbeitet wurde, sind in Tabelle 3 aufgeführt.

Tabelle 3: Liste der benutzten Kits.

Verwendung	Produkt	Hersteller
Aufreinigung von DNA-Fragmenten	QIAquick Gel Extraction Kit	Qiagen
	QIAquick PCR Purification Kit	
DNase Verdau	innuPREP DNase I Digest Kit	Analytik Jena
Größenbestimmung von DNA-Fragmenten	GeneRuler DNA Ladder Mix	MBI Fermentas
	MassRuler DNA Ladder Mix	
Microarray	ULS Fluorescent Labeling Kit for Agilent Arrays (with Cy3 and Cy5)	Kreatech
	Gene Expression Hybridization Kit	Agilent
	A-GEOD-13998 - Agilent-028409	Agilent
	Bacillus megaterium Bm3_100512	
Nukleinsäurefarbstoff	GelStar™ Nucleic Acid Gel Stain	Lonza
Plasmidisolation	QIAprep Spin Miniprep Kit	Qiagen

RIN-Wert Bestimmung	RNA 6000 Nano Kit	Agilent
RNA Präparation	innuPREP RNA Mini Kit	Analytik Jena

2.4 Bakterienstämme und Plasmide

Die verwendeten Stämme und Plasmide, sowie die aus der Arbeit resultierenden Plasmide sind Tabelle 4 zu entnehmen.

Tabelle 4: Stämme und Plasmide, die Bestandteil der Arbeit waren.

Bezeichnung	Charakterisierung	Referenz / Quelle
<i>Escherichia coli</i> Stamm		
DH10B	<i>F</i> - mcrA Δ (mrr-hsdRMS-mcrBC) ϕ 80dlacZ Δ M15 Δ lacX74 deoR recA1 endA1 araD139 Δ (ara, leu) 7697 galU galK λ - rpsL nupG	Gibco Life Technologies
<i>Bacillus subtilis</i> Stamm		
DSM 402	<i>B. subtilis</i> subsp. enthält das Plasmid pBC16 (DSM-Nr .: 23521)	DSMZ, Braunschweig
<i>Bacillus megaterium</i> Stämme		
DSM 319	Wildtyp	DSMZ, Braunschweig
WH320	<i>lacZ</i> -Mutante von DSM 319	Rygus <i>et al.</i> , 1991
WH323	<i>xylA1-spoVG-lacZ</i> -Mutante von WH320	Rygus and Hillen, 1992
Plasmide		
pRBBm99	pSSBm85-Derivat, Amp ^r , Tet ^r ohne <i>xylR</i> und P _{<i>xylA</i>} - <i>gfp</i>	Münch <i>et al.</i> , 2015
pMM1520	pSSBm85-Derivat, Amp ^r , Tet ^r ohne P _{<i>xylA</i>} - <i>gfp</i> mit <i>xylR</i>	Malten <i>et al.</i> , 2005
pJS72	pETDuet-Derivat, Amp ^r <i>Strep-tag-mCherry</i> - <i>mreBH</i>	Soufo <i>et al.</i> , 2010
pSSBm85	Promotoroptimierter Shuttle-Vektor für induzierbare eGFP-Produktion mit Xylose;	Stammen <i>et al.</i> , 2010

	Amp ^r , Tet ^r , <i>xylR</i> , P _{<i>xylA</i>} - <i>gfp</i>	
pBC16	natürlich vorkommendes Plasmid aus <i>Bacillus cereus</i> , Vorläufer für pSSBm85	Bernhard <i>et al.</i> , 1978
pKMMBm1	eingeführte <i>Xma</i> I-Stelle am 5'-Ende von <i>xylR</i> ersetzt <i>xylR</i> -Stoppcodon mit einer <i>Kpn</i> I-Stelle in pSSBm85	Münch <i>et al.</i> , 2015
pKMMBm2	klonieren von mCherry in pKMMBm1 über <i>Kpn</i> I und <i>Afl</i> III; <i>xylR</i> -mCherry-Fusion	Münch <i>et al.</i> , 2015
pKMMBm5	pSSBm85-Derivat, Amp ^r , Tet ^r , <i>xylR</i> -Mutante, P _{<i>xylA</i>} - <i>gfp</i>	diese Arbeit
pKMMBm6	pKMMBm2-Derivat, Amp ^r , Tet ^r , <i>xylR</i> -mCherry, P _{<i>xylA</i>} - <i>gfp</i> , multiple <i>xylR</i> -Bindestellen	diese Arbeit
pKMMBm7	pKMMBm2-Derivat ohne <i>gfp</i> , Amp ^r , Tet ^r , <i>xylR</i> -mCherry, P _{<i>xylA</i>}	diese Arbeit

2.5 Mikrobiologische Methoden

2.5.1 Kulturmedium für *Escherichia coli*

Die Kultivierung in *E. coli* wurde im Komplexmedium LB durchgeführt. Das Medium wurde wie folgt hergestellt:

LB-Medium:	Trypton	10 g
	Hefeextrakt	5 g
	NaCl	5 g
	H ₂ O _{deion}	ad 1000 ml

Für die Herstellung von LB-Festmedium wurde 15 g/l Agar verwendet.

2.5.2 Kulturmedium für *Bacillus megaterium*

B. megaterium wurde in einem halbdefinierten A5-Medium (Malten *et al.*, 2005) kultiviert. Um eine vorzeitige Induktion des Plasmidsystems zu verhindern, wurde anstatt Glucose das Monosaccharid Fructose verwendet. Die Zusammensetzung ist wie folgt:

A5-Medium:	10 x Pufferstammlösung	100 ml
	8 x Grundlösung	100 ml
	30 % Fructose	100 ml
	Spurenelemente	1 ml
	Eisensulfat	1 ml
	H ₂ O _{deion}	ad 1000 ml

Die Bestandteile der einzelnen Komponenten können der Tabelle 5 entnommen werden. Zur Herstellung von Festmedium wurde 15 g/l Agar zugegeben.

Tabelle 5: Zusammensetzung der Lösungen für das Medium A5.

Lösung	Reagenz	Menge
10 x Pufferstammlösung	KH ₂ PO ₄	3,52 g
	NaH ₂ PO ₄ x 2 H ₂ O	7,26 g
	ad 100 ml H ₂ O _{deion}	
	autoklavieren	
8 x Grundlösung	(NH ₄) ₂ SO ₄	8,00 g
	MgSO ₄ x 7 H ₂ O	1,20 g
	Hefeextrakt	2,00 g
	ad 800 ml H ₂ O _{deion}	
	autoklavieren	
30% Fructose	Fructose	30,00 g
	ad 100 ml H ₂ O _{deion}	
	sterilfiltrieren	

Spurenelemente	MnCl ₂ x 4 H ₂ O	4,00 g
	CaCl ₂ x 2 H ₂ O	5,30 g
	(NH ₄)Mo ₇ O ₂₄ x 4 H ₂ O	0,25 g
	CoCl ₂ x 6 H ₂ O	0,25 g
	ad 100 ml H ₂ O _{deion} sterilfiltrieren	
Eisensulfat	FeSO ₄ x 7 H ₂ O	0,25 g
	ad 100 ml H ₂ O _{deion} sterilfiltrieren	

2.5.3 Antibiotika

Diese wurden als konzentrierte Stocklösung hergestellt, sterilfiltriert und zum abgekühlten Medium pipettiert. Lösungsmittel und Konzentrationen können Tabelle 6 entnommen werden.

Tabelle 6: Antibiotika und ihre Konzentrationen.

Antibiotikum	Solvatation	Stocklösung	Endkonzentration	Stamm
Carbenicillin	H ₂ O _{deion}	100 mg/ml	100 µg/l	<i>E. coli</i>
Tetracyclin	70 % EtOH (v/v)	10 mg/ml	10 µg/ml	<i>B. megaterium</i>

2.5.4 Allgemeine Lösungen

Lösungen die im Laboralltag wichtig waren, werden in Tabelle 7 aufgezeigt.

Tabelle 7: Allgemeine Lösungen, die für verschiedene Experimente gebraucht wurden.

Lösung	Reagenz	Menge
10 x PBS pH 7,4	NaCl	150 mM
	KCl	3 mM
	NaPO ₄ -Puffer pH 7,4	10 mM

1 x TE	Tris-HCl pH 7,5	10 mM
	EDTA pH 8,0	1 mM

2.5.5 Lösungen für die Herstellung und Transformation von Protoplasten

Lösungen, die zur Herstellung und Transformation von Protoplasten benötigt wurden, können Tabelle 8 entnommen werden.

Tabelle 8: Lösungen, die der Herstellung und Transformation von Protoplasten dienen.

Lösung	Reagenz	Menge
2 x AB3	AB3 (Difco)	35 g
	H ₂ O _{deion}	ad 1000 ml
	für 15 min autoklavieren	
PEG-P	PEG-6000	400 g
	2 x SMM	ad 1000 ml
	für 15 min autoklavieren	
2 x SMM	Maleinsäure	40 mM
	NaOH	80 mM
	MgCl ₂ x 6 H ₂ O	40 mM
	Saccharose	1 M
	in angegebener Reihenfolge	
	in H ₂ O _{deion} lösen	
	pH 6,5 mit HCl einstellen	
	sterilfiltrieren	
SMMP-Lösung	2 x AB3	
	2 x SMM	
	zu gleichen Teilen mischen	

Lösung A	Saccharose	600 mM
	MOPS	60 mM
	NaOH	30 mM
	pH 7,3 mit NaOH einstellen	
	sterilfiltrieren	
Lösung B	Agar	14 g/l
	Casamino-Acids	0,7 g/l
	Hefe-Extrakt	35 g/l
	in H ₂ O _{deion} lösen	
	für 15 min autoklavieren	
8 x CR5-Salze	K ₂ SO ₄	11 mM
	MgCl ₂ x 6 H ₂ O	394 mM
	KH ₂ PO ₄	3 mM
	CaCl ₂	159 mM
	für 15 min autoklavieren	
Prolin	L-Prolin	12 % (w/v)
	in H ₂ O _{deion} lösen	
	sterilfiltrieren	
Glucose	D-Glucose	20 % (w/v)
	in H ₂ O _{deion} lösen	
	sterilfiltrieren	
CR5-Topagar	Lösung A	1,25 ml
	Lösung B	713 µl
	CR5-Salze	288 µl
	Prolin	125 µl
	Glucose	125 µl

2.5.6 Lösungen für FISH

Lösungen, die für FISH-Experimente hergestellt wurden, sind in Tabelle 9 dargestellt.

Tabelle 9: Lösungen zur Durchführung der FISH-Experimente.

Lösung	Reagenz	Menge
GTE	Glucose	50 mM
	EDTA	10 mM
	Tris pH 7,5	20 mM
NaPO ₄ -Puffer pH 7,4	1 M Na ₂ HPO ₄	77,4 ml
	1 M NaH ₂ PO ₄	22,6 ml
20 x SSC	Tri-Natriumcitrat pH 7,0	15 mM
	NaCl	150 mM
Lösung zum Blocken	Formamid	70 %
	SSC	2 x
	Salmon sperm DNA	1 mg/ml

2.6 Zellbiologische Methoden

2.6.1 Sterilisation

Sofern möglich, wurden sämtliche Lösungen, Medien und dessen Bestandteile bei 1 bar Überdruck für 20 min. bei 121 °C autoklaviert. Temperaturempfindliche Reagenzien wurden sterilfiltriert (Porengröße 0,2 µM).

2.6.2 Kultivierung von Bakterien

Zur Kultivierung wurden die Stämme auf LB-Platten und die plasmidtragenden Stämme

auf A5-Platten (Malten *et al.*, 2006) mit Antibiotikum (Kap. 2.5.3) aus dem Glycerolstock angeimpft. Nach dem Bebrüten bei 37 °C o/n wurden die plasmidtragenden Stämme auf eine neue Platte überimpft und nochmals bei der selben Temperatur über Nacht bebrütet. Um die Vitalität der Bakterienkultur zu gewährleisten, wurden zum Animpfen einer Vorkultur in Flüssigmedium ausschließlich frisch inokulierte und bewachsene Platten verwendet. Die Inkubation der Vorkultur für die Hauptkultur erfolgte für 4 bis 5 Std bei 220 rpm und 37 °C im Wasserbadschüttler. Die Vorkultur wurde in der Nacht mit Hilfe einer Zeitschaltuhr gestartet, um die Fitness der Bakterien am nächsten Morgen zu garantieren. Demzufolge befanden sich die Kulturen in der späten exponentiellen Phase. Die mittels Photometer bestimmte OD bei einer Wellenlänge von 578 nm wurde in 50 ml A5-Medium auf eine optische Dichte von 0,05 verdünnt und im Schikanekolben bis zur gewünschten OD bzw. Probeentnahme inokuliert.

Zum Messen der Produktionsraten von eGFP und mCherry sowie zur online-Detektion des Zellwachstums diverser Kulturen im Vergleich, wurde eine 48-well-FlowerPlate mit 1 ml Bakteriensuspension ($OD_{578} = 0,05$) versetzt. Diese wurde mit einer gaspermeablen Folie verschlossen und zum Schütteln bei 1400 rpm o/n in einem Mikrobioreaktor (BioLector) gestellt. Um die Streulichtwerte in eine optische Dichte umzuwandeln, war eine Kalibrierungsgerade erforderlich.

2.6.3 Induktion der plasmidtragenden Stämme

Die Induktion der plasmidtragenden Stämme und somit Produktion des Reportergens eGFP bzw. der Reportergene eGFP und mCherry erfolgte mit 0,5 % Xylose bei einer OD_{578} von 0,2. Die Kontrollen wurden auf dieselbe Weise induziert, auch wenn sie keine Plasmide besaßen. Im Vergleich wurde stets die Kultur vor und nach Induktion betrachtet.

2.6.4 Konservierung von Bakterien

Von einer frischen Bakterienkultur, die über Nacht im entsprechenden Medium gewachsen ist, wurde 700 µl Suspension entnommen und zu 300 µl 50 %igem Glycerin pipettiert. Die Lagerung erfolgte bei -80 °C.

2.7 Molekularbiologische Methoden

2.7.1 Zellzahlbestimmung und Ermitteln von Wachstumskurven

Die Zelldichte wurde zum einen optisch durch Absorption mittels Photometer bei 578 nm und zum anderen mit Hilfe des BioLectors über Streulicht bestimmt. Erhielt man bei der Absorption der photometrisch ermittelten Werte eine Zahl ≥ 1 , wurde die Probe verdünnt.

Wachstumskurven wurden von allen zu untersuchenden plasmidtragenden Stämmen einschließlich ihrer Kontrollen erstellt. Die Vorgehensweise der Kultivierung wurde in Kap. 2.6.2 beschrieben. Sofern möglich, wurden technische und biologische Replikate angefertigt. Die Messung wurde entsprechend der Wachstumsphase alle 10 bis 30 min durchgeführt. Diese Zeitintervalle wurden analog im BioLector verwendet.

2.7.2 Extraktion von Nukleinsäuren

DNA wurde mit Hilfe des QIAprep Spin Miniprep Kits von Qiagen extrahiert und RNA über das innuPREP Mini Kit von Analytik Jena. Zum Lysieren der Zellen wurde der Lysepuffer mit Lysozym (50 µg pro Probe) versetzt. Die Behandlung erfolgte für 15 min bei 37 °C unter leichtem Schütteln.

Generell wurden die Nukleinsäuren in H₂O_{deion} resuspendiert. Zum Lösen von RNA wurde dieses mit DEPC behandelt. Dafür wurde zu 1 Liter H₂O_{deion} 1 ml DEPC (handelsübliche Lösung der Firma Sigma) zugesetzt und gut vermischt. Die Einwirkzeit betrug mindestens eine Stunde, wenn möglich auch über Nacht, wobei des Öffern

erneut gut durchgemischt wurde. Nach dem Autoklavieren zerfällt die gesundheits-schädliche Substanz zu Ethanol und CO₂.

Jeweils 2 µl der Nukleinsäuren wurden im NanoDrop vermessen. Dieser misst das Verhältnis A260/A280 nm. Zur vorläufigen Beurteilung der Reinheit ist ein Verhältnis zwischen 1,8 und 2,0 wichtig. Eine erhöhte Extinktion bei 230 kann ebenfalls auf eine Verunreinigung hindeuten, so dass das Verhältnis A260/A230 nm über 2,0 liegen sollte. Zur sicheren Bestimmung der Reinheit von RNA wurde der Bioanalyzer von Agilent verwendet. Die Qualifizierung erfolgte mit dem RNA 6000 Nano Kit. Dabei wird die intakte ribosomale RNA (5S, 16S und 23S) und ihre bereits fragmentierten Bestandteile – sofern vorhanden – mittels Kapillarelektrophorese der Größe nach aufgetrennt und sowohl über ein Elektropherogramm, als auch über ein virtuelles Gelbild visualisiert. Das Verhältnis der 23S zu 16S rRNA wird dabei auf einer Zahlenskala von 1 bis 10 ermittelt. Der Wert 10 steht dabei für die höchste Qualität, hingegen bedeutet der Zahlenwert 1 komplett degradierte RNA (Schroeder *et al.*, 2006). Durch die neuen MIQE-Richtlinien (Bustin *et al.*, 2009) müssen für die qRT-PCR RIN-Werte über 5,0 erzielt werden (Fleige and Pfaffl, 2006). Für den Microarray lag die Grenze bei Werten $\geq 8,0$.

2.7.3 Herstellung chemisch kompetenter Zellen für die Transformation in *E. coli*

Zur Herstellung chemisch kompetenter Zellen für die Transformation in *E. coli* wurde der Stamm DH10B in LB-Medium ohne Antibiotikum angeimpft und über Nacht bei 37 °C und 200 rpm inokuliert. Von dieser Vorkultur wurden 200 µl entnommen und in einen EMK (Nennvolumina 1000 ml) zu 200 ml LB-Medium gegeben. Demzufolge wurde mit $\frac{1}{1000}$ Volumen angeimpft. Die Zellen wurden bis zu einer OD₅₇₈ = 0,5 inkubiert und 10 min zum Abkühlen auf Eiswasser gestellt. Anschließend wurden die Zellen für 5 min bei 3000 rpm und 4 °C zentrifugiert. Nachdem das Medium dekantiert wurde, wurden die Zellen in 50 ml eiskaltem 0,1 M MgCl₂ gelöst und erneut für 5 min bei 2500 rpm und 4 °C sedimentiert. Das Pellet wurde in 50 ml eiskaltem 0,1 M CaCl₂ resuspendiert und für 20 min ins Eiswasser gestellt. Zuletzt wurden die Zellen nochmals 5 min bei 2000 rpm und 4 °C zentrifugiert und in 10 ml eiskalter 0,1 M CaCl₂-Lösung,

die 10 % Glycerol enthielt, resuspendiert. Aliquots von 50 µl wurden in ein Eppendorfröhrchen (Nennvolumina 1,5 ml) überführt und sofort in flüssigem Stickstoff schockgefroren. Die Lagerung erfolgte bei -80 °C. Die Transformationsrate τ lag zwischen 10^6 und 10^7 Zellen / µg DNA.

2.7.4 Herstellung kompetenter Zellen für die Transformation in *B. megaterium*

Einer 50 ml Übernachtskultur des Wildtypstammes DSM 319 wurde 1 ml entommen und in 50 ml neues LB-Medium transferiert. Die Kultur im EMK (Nennvolumina 250 ml) wurde bei 200 rpm und 37 °C inkubiert, bis eine $OD_{578} = 1,0$ erreicht war. Im weiteren Verlauf wurde die Suspension 15 min bei 5000 rpm und 4 °C abzentrifugiert und das Pellet in frisch hergestellter SMMP-Lösung, die 1 µg/ml Lysozym enthielt, resuspendiert. Die darauffolgende Protoplastierung dauert unter leichtem Schütteln bei 37 °C ca. 30 min und wurde mit Hilfe des Mikroskops überprüft. Dabei sollten überwiegend runde Zellen zu sehen sein, da die Mureinschicht der Zellwand aufgelöst wurde. Anschließend wurden die Zellen für 10 min bei 3000 rpm und RT sedimentiert. Das Pellet wurde vorsichtig in 5 ml SMMP gelöst, erneut mit den selben Kriterien zentrifugiert und ein weiteres Mal in 5 ml SMMP resuspendiert. Die Protoplastensuspension wurde vorsichtig mit 750 µl Glycerin (87 %) vermenget und zu jeweils 500 µl portioniert. Die Aufbewahrung erfolgte bei -80 °C.

2.7.5 Transformation in *E. coli*

Die Transformation in *E. coli* wurde mit chemisch kompetenten DH10B Zellen durchgeführt. Diese wurden für 15 min auf Eis aufgetaut. Dann pipettierte man zwischen 10 bis 50 ng Plasmid-DNA zu 50 µl chemisch kompetenten Zellen und inkubierte diese für 15 min auf Eis. Nun erfolgte ein Hitzeschock von 45 sek bei 42 °C im Thermomixer. Die Transformation wurde für 2 min auf Eis gestellt. Abschließend wurden 200 µl LB-Medium ohne Zusatz von Antibiotikum zugegeben und sofort bei 37 °C unter leichtem Schütteln für 1 Std inkubiert.

Je nach Transformationsrate der kompetenten Zellen wurden entsprechende Aliquots auf LB-Platten mit Antibiotikum ausplattiert.

2.7.6 Transformation in *B. megaterium*

Alle Lösungen wurden hierfür auf Raumtemperatur temperiert. Zwischen 1 bis 5 µg Plasmid-DNA wurde in H₂O_{deion} vorgelegt. In Zentrifugenröhrchen (Nennvolumina 15 ml) wurde 1,5 ml PEG-P vorgegeben. Je 500 µl der aliquotierten Protoplastensuspension wurde mit der DNA vermennt und vorsichtig in das Falconröhrchen zum PEG-P gemischt. Nach zweiminütiger Inkubation bei Raumtemperatur wurden 5 ml SMMP zupipettiert und abermals vorsichtig gemischt. Der Ansatz wurde für 10 Minuten bei 3000 rpm und RT abzentrifugiert. Der Überstand wurde sofort dekantiert und das Pellet in 500 µl SMMP ohne zu vortexen resuspendiert. Die Suspension wurde in ein Eppendorfröhrchen mit 1,5 ml Nennvolumina überführt. Bei 30 °C wurde der Transformationsansatz zunächst für 45 min ohne Schütteln inkubiert, danach für weitere 45 min bei 300 rpm. Gegen Ende der Inkubationszeit wurde der Ansatz zu 2,5 ml vorgewärmtem CR5-Topagar gegeben, der sich in Falconröhrchen befand und durch Rollen vorsichtig vermennt, aber nicht umgedreht wurde. Der Inhalt des Falcons wurde zügig auf vorgewärmte tetracyclinhaltige LB-Platten verteilt und o/n bei 30 °C inkubiert.

Von den gewachsenen und überprüften Kolonien wurden Glycerolstocks angelegt.

2.7.7 Plasmidkonstrukte

Alle in dieser Studie verwendeten Plasmide sind in der Tabelle 4 dargestellt. Die angefertigten Vektoren im Stamm DSM 319 basierten auf dem Plasmid pSSBm85 (Abb. 9C). Bei den verwendeten Kontrollen handelte es sich entweder um Derivate (Leerplasmide) oder dem Wildtypstamm. Das Vorgängerplasmid pWH1508 wurde mit pBR327 und pBC16 (Abb. 9A) fusioniert (Rygus *et al.*, 1991). Das Plasmid trägt das Gen *gfp*⁺, eine Variante des Wildtyp-GFP, dessen Genprodukt (eGFP) eine erhöhte

Fluoreszenz aufweist (Scholz *et al.*, 2000). Ursprünglich wurde diese GFP-Variante für das pMUTIN-GFP + Plasmid (Kaltwasser *et al.*, 2002) beschrieben.

Für die Gestaltung einer C-terminalen XylR-mCherry-Fusion im Expressionssystem wurde ausgehend vom Produktionsplasmid pSSBm85 (Stammen *et al.*, 2010) das Stopcodon von *xylR* durch eine neue *KpnI*-Schnittstelle ersetzt. Dies wurde durch die Einführung eines synthetischen Fragments, das das gesamte *xylR*-Gen inklusive der beiden Schnittstellen *AflIII* und *SpeI* beinhaltete bewerkstelligt, was zu pKMMBm1 führte. Als Primer wurden die Oligonukleotide *xylR*-mCherry forward (5'-TATCAGGTACCGTGAGCAAGGGCGAGGAG-3') und *xylR*-mCherry reverse (5'-TATCACTTAAGTTACTTGTACAGCTCGTCCATG-3') verwendet, um das Gen das mCherry codiert aus dem Plasmid pJS72 zu amplifizieren und dabei zusätzlich die flankierenden Restriktionsstellen *AflIII* und *KpnI* einzuführen. Dieses Amplikon wurde in den mit *AflIII* und *KpnI* geschnittenen Vektor pKMMBm1 kloniert. Die daraus resultierende XylR-mCherry-Fusion entstand durch die Klonierung des mCherry codierten *AflIII/KpnI*-Fragments stromabwärts (downstream) von *xylR* und wurde als pKMMBm2 bezeichnet (Abb. 8 und Abb. 9D). Die verbleibenden potenziellen Membranbindestellen BA1 und BA2 (Korn *et al.*, 1983 sowie Tanaka and Sueoka, 1983) sind orange markiert.

Das Δ *xylR*-Mutationsplasmid pKMMBm5 (Abb. 9E) wurde durch Restriktion mit *XmaI* und *KpnI* aus pKMMBm1 generiert. Dabei wurde das *xylR*-Gen entfernt und die verbliebenen klebrigen Enden (sticky ends) mit Klenow-Polymerase in glatte Enden (blunt ends) umgewandelt. Die darauf folgende Religation erzeugte diesen Vektor.

Das Plasmid pKMMBm6 (Abb. 9F) basiert auf pKMMBm2 und besteht aus einem synthetischen Fragment von zehn sich wiederholenden Bindestellen für XylR (Rygas *et al.*, 1991). Die Sequenz der Tandem Repeats lautet:

5'-TTGAGTTAGTTTGTGTTTATCCAATAAACTAACTTT-3'. Dazwischen befinden sich Zufallssequenzen von jeweils zehn Basen, die nicht als Restriktionsschnittstelle charakterisiert werden konnten. Die Sequenz, die mittels *EagI* eingefügt wurde, ist wie folgt:

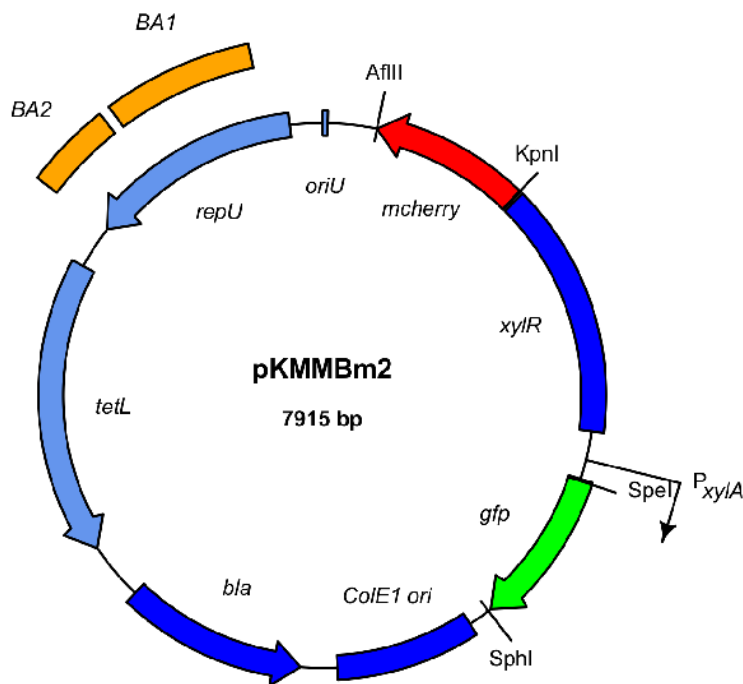


Abbildung 8: Plasmidkarte von pKMMBm2. Dieses Plasmid basiert auf den Vektor pSSBm85, enthält aber zusätzlich eine translationale mCherry-Fusion am C-Terminus mit dem Repressor XylR. Die dabei verbleibenden potenziellen Membranbindestellen (BA1 und BA2) sind orange markiert.

CGGCCGGGTAAGCTTGAGTTAGTTTGTTTATCCAATAAACTAACTTTTAG
AACCAGTTTGAGTTAGTTTGTTTATCCAATAAACTAACTTTAGATGTAGACTT
GAGTTAGTTTGTTTATCCAATAAACTAACTTTGTCTAACGCTTTGAGTTAGTTT
GTTTATCCAATAAACTAACTTTGTATGCATCGTTGAGTTAGTTTGTTTATCCAAT
AACTAACTTTGTTTCAGCATATTGAGTTAGTTTGTTTATCCAATAAACTAACTT
TTGTACAAGTCTTGAGTTAGTTTGTTTATCCAATAAACTAACTTTAGACGATTC
GTTGAGTTAGTTTGTTTATCCAATAAACTAACTTTTGAGACTTGATTGAGTTAG
TTTGTTTATCCAATAAACTAACTTTCAGTTCGTACTTGAGTTAGTTTGTTTATC
CAATAAACTAACTTTGCCTGCTGAGGCGGCCGC.

Das letzte Plasmid pKMMBm7 basiert ebenfalls auf pKMMBm2, dem das Gen für eGFP über die Restriktionsendonucleasen *NotI* und *SpeI* entfernt wurde. *Sticky ends* wurden mit Klenow aufgefüllt und mittels Ligase religiert.

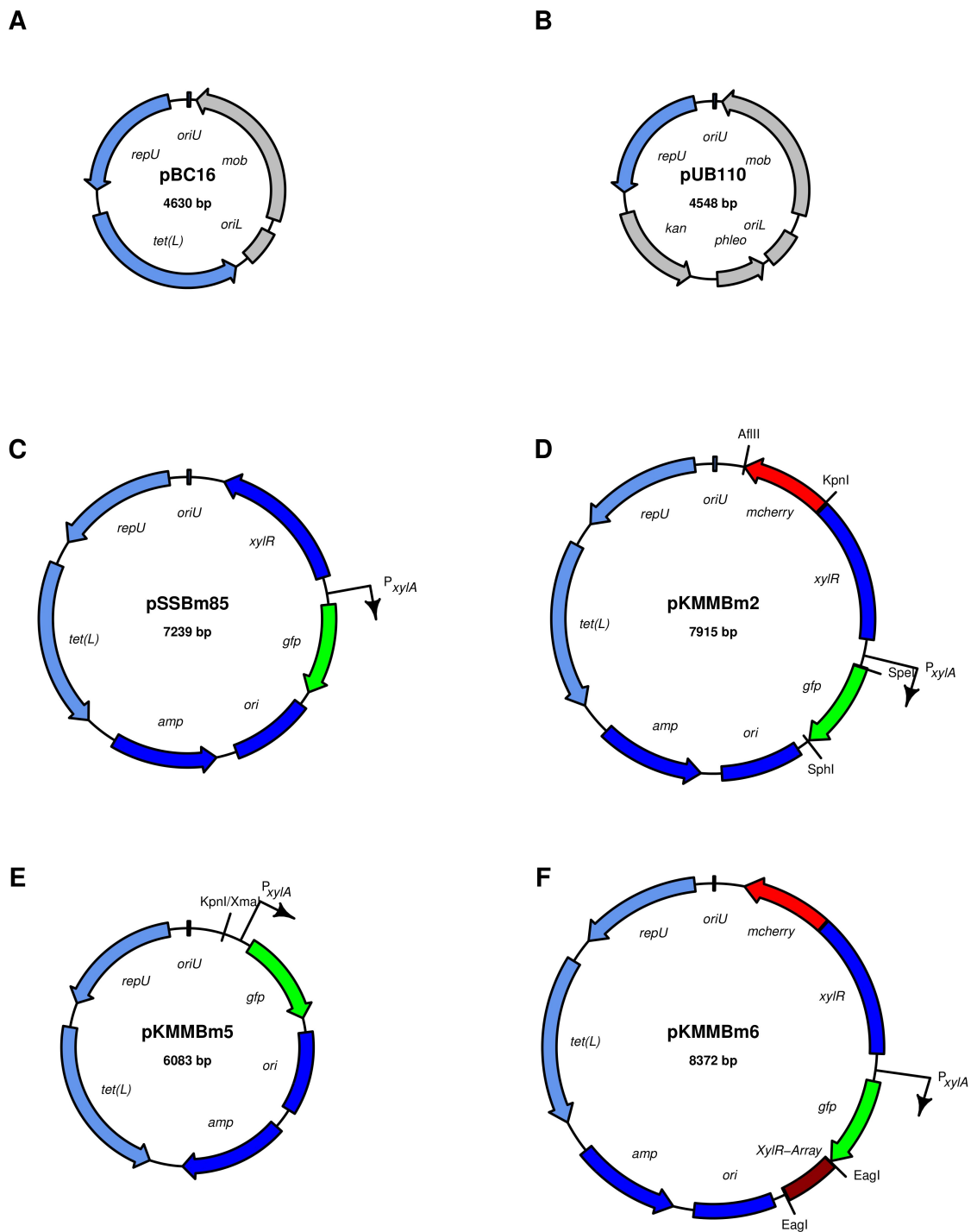


Abbildung 9: Wichtige Vektoren im Überblick. A) Das Plasmid pBC16 wurde mit pBR327 fusioniert, wodurch das Vorgängerplasmid pWH1508 entstand. B) Der Vektor pUB110 ist bis auf das Resistenzgen identisch zum Plasmid pBC16 (A). Er enthält vier nachgewiesene Membranbindestellen. C) Bei pSSBm85 handelt es sich um das Produktionsplasmid, das mit Einführen einer C-terminalen XylR-mCherry-Fusion in pKMMBm2 (D) resultierte. E) Das $\Delta xylR$ -Mutationsplasmid pKMMBm5 entstand durch Herausschneiden des *xylR*-Gens. F) Basierend auf pKMMBm2 (D) wurde pKMMBm6 konstruiert. Dieser Vektor besteht aus zehn, sich wiederholenden Bindestellen für XylR.

2.7.8 Detektion der Fluoreszenz

Die *in vivo* Fluoreszenz wurde in den Zellkulturen mit einem Spectrofluorimeter von Jasco und dem BioLector von m2p-labs detektiert. Zellen wurden wie unter 2.6.2 beschrieben kultiviert, Kontrollen wurden nicht mit Xylose behandelt und Proben dementsprechend mit 0,5 % Xylose induziert. Mittels Replikaten wurde das Wachstum parallel entweder über das Photometer oder über die Streulichtmessung im BioLector gemessen. Die Anregung wurde bei einer Wellenlänge von 488 nm für eGFP und 587 nm für mCherry vorgenommen. Die Emission fand für eGFP bei 507 nm und für mCherry bei 610 nm statt.

2.7.9 DNA-Microarray

DNA-Arrays spielen eine bedeutende Rolle bei der Analyse der differentiellen Genexpression. Das Trägermaterial kann mit unterschiedlichen DNA-Fragmenten, wie z. B. cDNA, Fragmente von PCR-Produkten und Oligonukleotiden beschichtet werden. Bei dem in dieser Arbeit verwendeten Array handelt es sich um einen Oligonukleotid-Array von Agilent (A-GEOD-13998 - Agilent-028409 *Bacillus megaterium* Bm3_100512). Wird die zu untersuchende mRNA mit Fluoreszenzfarbstoffen markiert, kann die Sonde komplementär an die einsträngig angeordnete (array = Anordnung) DNA hybridisieren. Dabei zeigen sich Änderungen von Genexpressionen in Abhängigkeit von verschiedenen Bedingungen. Dies kann mit Hilfe der Fluoreszenz über unterschiedliches Markieren der verschiedenen Zustände gemessen werden. Aus den beiden Fluoreszenzfarbstoffen entsteht eine Mischfarbe, die dann mit einer Laserkamera detektiert werden kann.

Für diese Zweifarbendetektion wurden Cyanine eingesetzt, die im Fall von Cy3 im Grünbereich bei 550 nm und von Cy5 im Rotbereich bei 650/670 nm fluoreszieren. Bei einer Wellenlänge von 570 nm emittiert Cy3, wohingegen Cy5 bei 649 nm emittiert. Im Verlauf des Experiments wird eine Probe mit Cy3, die andere mit Cy5 markiert. Der Sondenpunkt erscheint gelb, wenn die Expression der beiden Proben in etwa gleich groß ist. Sobald die Expression eines Transkripts überwiegt, leuchtet die Sonde an dieser

Stelle entsprechend grün oder rot auf. Da die Sensitivität des Labelns und des Hybridisierens unterschiedlich ist, muss eine Probe reziprok markiert werden. Dieses Verfahren wird als Dye Swap bezeichnet.

Zum Ermitteln der Expression musste ein Grenzwert festgelegt werden der besagt, ab wann eine signifikante Regulation vorliegt. Dieser Wert wurde auf 2,0 festgelegt (Ehrenreich, 2006).

In dieser Arbeit wurden pKMMBm2-Kulturen vor und nach Induktion mit Xylose verglichen. Die Zellen wurden wie in Kap. 2.6.2 und 2.6.3 kultiviert und induziert. Für diesen Versuch wurden vier biologische Replikate verwendet. Die RNA wurde wie in Kap. 2.7.2 beschrieben isoliert und die RIN-Werte mittels Bioanalyzer (Agilent) bestimmt. Die RNA wurde nach Schritt 2 des “Two-Color Microarray-Based Gene Expression Analysis Protocol“ von Agilent Technologies hybridisiert und fragmentiert. Die Markierung mit Cy3 und Cy5 wurde mit dem ULS Fluorescent Labeling Kit von Kretech vorgenommen. Bei jeweils drei Replikaten wurde die Probe vor Induktion mit Cy5 markiert und die nach Induktion mit Cy3. Das jeweils vierte Replikat wurde reziprok markiert. Die Einbaurate wurde anhand des UV-VIS Spektrophotometers von peqlab ermittelt. Jeweils 300 ng markierter RNA wurde mit dem Gene Expression Hybridization Kit von Agilent hybridisiert und im Microarray-Scanner (Agilent) analysiert. Dazu wurden die markierten Proben mit 5 µl Blocking Agent (10 x) versetzt und auf ein Endvolumen von 24 µl aufgefüllt. Nach dem Zufügen von 1 µl Fragmentierungspuffer (25 x) wurde der Ansatz für 30 min bei 60° C inkubiert. Im Anschluss erfolgte ein Zugabe von 25 µl Hybridisierungspuffer HI-RPM (2 x) und ein Abkühlen auf Eis. Die Probe wurde luftblasenfrei auf ein Gasket Slide (8 x 15 K) aufgetragen, der Array Slide aufgelegt und 17 Std bei 65° C sowie 10 rpm im Hybridisierungsofen inkubiert. Der Array Slide wurde im Waschpuffer 1 vom Gasket Slide gelöst und in temperierter Waschlösung 2 (37° C) für 1 min inkubiert. Der fertige Microarray wurde mit dem High-Resolution Microarray Scanner von Agilent Technologies mittels der Software Agilent Scan Control (Version 8.4.1) eingelesen und die Daten anhand der Software Feature extraction (Version 10.7.3.1) extrahiert. Die statistische Auswertung der Microarrays wird in Kapitel 2.8.2 beschrieben.

2.7.10 Quantitative Real-Time PCR (qRT-PCR)

Die Methode der Echtzeit-PCR bietet unter Verwendung einer Fluoreszenzmessung die Möglichkeit zur Quantifizierung von Genexpressionen. Extrahierte mRNA wird zunächst in cDNA umgeschrieben und dann mittels PCR amplifiziert. Während die Methode des DNA-Microarray ein globales Bild zeigt, vermittelt die qRT-PCR über spezifische Primer eine Expression einzelner ausgewählter Gene. Der Reaktion beigefügte Fluoreszenzfarbstoffe ermöglichen eine Quantifizierung des PCR-Produkts (Amplikon). Die gemessene Fluoreszenz nimmt proportional zum entstehenden PCR-Produkt exponentiell zu.

Ein wichtiger Bestandteil der Messung ist die Bestimmung des C_q -Werts. Er gibt an, wann die Fluoreszenz erstmalig signifikant über den Hintergrund steigt. Im Zuge der Standardisierung von qRT-PCR-Daten entsprechend des RDML-Formats (Real-Time PCR Data Markup Language; Lefever *et al.*, 2009), wurden die bisher verwendeten Abkürzungen CP für Crossing Point, Ct für Cycle threshold und Take-Off Point (TOF) mit C_q für quantification Cycle ersetzt.

Ein weiterer wichtiger Faktor ist die Effizienzberechnung. Nur bei einer Effizienz von $E = 2$ (100 %) findet eine Verdopplung statt. Meistens liegt diese jedoch unter diesem Optimum und muss somit in die Berechnung mit einbezogen werden.

Gesamt-RNA wurde unter Verwendung des innuPREP RNA Mini Kits isoliert. Im Anschluss wurde ein DNaseI-Verdau mit dem innuPREP DNaseI Digest Kit auf der Säule durchgeführt. Beide Kits wurden von der Firma Analytik Jena bezogen. Die Reinheit der RNA wurde mit Hilfe des Bioanalyzers überprüft, wobei RIN-Werte größer 5 erforderlich waren. Für die reverse Transkription wurde zwischen 7 µg und 10 µg Gesamt-RNA eingesetzt.

Zur Synthese der cDNA wurde die Sekundärstruktur der RNA für 10 min bei 70 °C aufgeschmolzen und mit DNA-Hexameren von Invitrogen, die aus Zufallssequenzen bestanden, versetzt und mit DEPC-behandeltem Wasser auf 16 µl aufgefüllt. Pro µg RNA wurden ca. 100 ng Primer verwendet.

Die Primerhybridisierung fand für 2 min bei 25 °C statt. Dazu wurde dem Ansatz 6 µl 5 x 1st Strand Buffer (Invitrogen), 10 mM DTT (Invitrogen), 330 µM dNTPs (GE Healthcare), sowie 20 U RNase Inhibitor (Invitrogen), bezogen auf 30 µl Gesamt-

volumen, hinzugefügt. Das Gesamtvolumen beträgt derzeit 27 µl, da noch keine reverse Transkriptase zugesetzt wurde. Die Angabe der Konzentrationen beziehen sich aber auf 30 µl. Die Synthese der Erststrang-cDNA fand nach Zugabe von 600 U SuperScriptII (Invitrogen) für 120 min bei 42°C statt. Für die No Reverse Transcription Control (NRT-Kontrolle) wurde an dieser Stelle von jeder vorhandenen RNA-Probe eine cDNA-Synthese ohne reverse Transkriptase durchgeführt. Dieser Ansatz soll beweisen, dass keine DNA mehr vorhanden ist, die das Ergebnis der Echtzeit-PCR beeinflussen könnte. Die RNA wurde unter Verwendung von 0,2 M NaOH und 0,1 M EDTA destabilisiert, der Ansatz mit 0,2 M HEPES und 1,2 M Natriumacetat neutralisiert. Die cDNA wurde mit Hilfe des PCR-Purification Kits von Qiagen gereinigt. Damit diese an die Säule bindet, muss ein pH-Wert kleiner 7,0 gewährleistet sein. Falls notwendig, wurden die Proben mit 1 M HCl versetzt. Die Ausbeute an cDNA betrug dabei zwischen 8 µg und 12 µg. Im Fall der NRT-Kontrolle konnte nichts quantifiziert werden.

Die quantitative Echtzeit-PCR wurde mit einem Gerät der Firma BioRad (CFX96) ausgeführt. Dabei betrug die Konzentration der eingesetzten cDNA 5 ng pro Ansatz. Bei dem verwendeten Fluoreszenzfarbstoff handelte es sich um SsoFast EvaGreen von BioRad, der in Form eines Supermixes bereits alles enthielt, außer cDNA und Primer. Der Reaktionsansatz bestand aus 8 µl H₂O_{deion}, 1 µl cDNA, 1 µl Primermix (jeweils 2 µM forward und reverse Primer) sowie 10 µl SsoFast EvaGreen Supermix. Die Negativkontrolle (No Template Control = NTC) beinhaltete H₂O_{deion} anstelle der cDNA. In aller Regel wurden drei technische Replikate verwendet. Konnten nur Duplikate gemessen werden, war die Differenz der beiden C_q-Werte stets ≤ 0,5. Die benutzte PCR-Maschine erfordert keine passive Referenz, so dass eine ROX-Zugabe nicht erforderlich war, um die Ergebnisse von unterschiedlichen Mikrotiterplatten zu vergleichen. Allerdings wurde auch hier auf eine Abweichung der einzelnen C_q-Werte von ≤ 0,5 geachtet. Durch die Zugabe dieses passiven Referenzfarbstoffs zum Reaktionspuffer können störende Schwankungen der Fluoreszenz ausgeglichen werden.

Die PCR wurde in entsprechenden Mikrotiterplatten (96-Well Plates, BioRad) nach folgendem Programm durchgeführt:

Schritt	Temperatur	Zeit	Zyklenanzahl
1	98 °C	3 min 30 sek	1
2	98 °C	5 sek	40 (ab Schritt 2)
	59 °C	15 sek	
	60 °C	15 sek mit Fluoreszenzmessung	
3	98 °C	10 sek	1
4	65 °C bis 95 alle 5 sek 0,5 °C Anstieg mit °C Messung der Fluoreszenz für die Schmelzkurve		1

Sämtliche Primerpaare (Tabelle 10) wurden über das Programm primer3plus (Untergasser *et al.*, 2013) generiert. Diese wurden zunächst bezüglich ihres Temperatur-optimums in einer herkömmlichen PCR-Maschine überprüft. Für die Gewährleistung der Signifikanz waren C_q -Werte zwischen 15 und 35 erforderlich. Die C_q -Werte beider Kontrollen (NRT / NTC) sollen fünf C_q -Werte üben denen der Proben liegen. Anhand der Schmelzkurve konnte sichergestellt werden, dass nur ein PCR-Fragment amplifiziert wurde.

Die relative Quantifizierung geht auf das $\Delta\Delta C_t$ -Modell zurück (Livak and Schmittgen, 2001). Dieses wurde weiterentwickelt und die Genexpression unter Einberechnung der Effizienz korrigiert (Pfaffl, 2001). Die Rohdaten wurden durch ein log-logistisches Modell mit fünf Parametern (Spiess *et al.*, 2008) gefittet. Dieses basiert auf einem log-logistischen 14 Modell (Abb. 10) mit vier Parametern (Tichopad *et al.*, 2003). Der FDM gibt den Wendepunkt der Kurve an. Ausschlaggebend für die Effizienzberechnung dieses Modells ist die Annahme, dass es einen Zeitpunkt gibt, bei dem die Effizienz gleich ist (exponentielle Phase). Andere Modelle besagen, da es in lebenden Systemen eine Plateau-Phase gibt, kann zu keinem Zeitpunkt die Effizienz gleichgroß sein, sondern variiert ständig und hängt von der vorherigen Effizienz ab (Rutledge, 2008). Das Ende der exponentiellen Phase wird durch das Maximum der zweiten Ableitung beschrieben (SDM) und stellt den C_q -Wert dar. Das basierende sigmoide (S-förmige)

vier-Parameter Modell zeigt durch die Assymetrie der Kurve im Bereich der Plateau-Phase eine Ungenauigkeit, so dass es durch einen fünften Parameter optimiert wurde (Spiess *et al.*, 2008). Ein Vergleich der beiden Modelle ist in Abbildung 11 aufgeführt. Die Datenanalyse für die Berechnung der C_q -Werte, der relativen Expressionsverhältnisse sowie einhergehender Statistik wurde unter Verwendung des qpcR-Pakets (Ritz and Spiess, 2008) durchgeführt und wird unter Kapitel 2.8.3 beschrieben.

Eine Vorauswahl potenzieller Referenzgene wurde unter Zuhilfenahme der Microarraydaten getroffen. Diese wurden mittels geometrischer Normalisierung unter Benutzung der Software geNorm (Vandesompele *et al.*, 2002) validiert. Die Datenanalyse erfolgte mit dem Bioconductor (Gentleman *et al.*, 2004) R-Package SLqPCR (Kohl, 2007), das in Kapitel 2.8.3 näher erläutert wird.

Tabelle 10: Verwendete Primer für die qRT-PCR. HKG (Housekeeping Genes) beschreiben die Primer zum Evaluieren der Referenzgene, wohingegen GOI (Gene of Interest) die Zielgene beschreiben.

Name	Beschreibung	Primersequenz
HKG		
<i>gap1</i>	Glyceraldehyde-3-phosphat- dehydrogenase, Typ I	for: CTTCCATCGTCACAGGTTTCCTTAC rev: CCAAGGCACTGTCATTAGTTCTTCC
<i>gap2</i>	Glyceraldehyde-3-phosphat- dehydrogenase, Typ I	for: TACACGGTTAGAATATCCGCTCTCG rev: ACAGCGAAGAGCCACTAGTATCTGG
<i>gatB</i>	Glutamyl-tRNA-amido-transferase, Yqey-Domäne	for: GATCTTGACGACCTGCTTTCTCAA rev: TTAGGAACCGACCTAACTGAAGCAG
<i>gyrB</i>	DNA-Gyrase, B-Untereinheit	for: TACATGGTGTAGGTGCCTCAGTTGT rev: ACTTTTAAGTCAGCAGCCGGTACAC
<i>heli1</i>	SNF2 Helikase	for: GATGTAATCCATACGTCAGCTGTGC rev: CCGGCTTCCTTGATTTATAACTGG
<i>heli2</i>	SNF2 Helikase	for: GTCGAACCTGTACACGGTGACTTTT rev: GCTTCCGTAATAGGTCTGTTCATGC
<i>rpoB</i>	RNA Polymerase, beta-Untereinheit	for: GCGGACGAAGTAGTAAAAGGTGAGA rev: GGCATCCTCATAGTTGTAACCACTC
<i>aepY</i>	Phosphonopyruvat Decarboxylase	for: CGCTACAGGAAAAACAGGAAGAGAG

		rev: GCTGGACAGTAGTATCCGTTGGTTG
<i>argR</i>	Arginin-Repressor	for: GGACGGTCGATACAAATACAGCTTA
		rev: GCGTATCTTCAGGAGTACGGCAAAT
<i>comEA</i>	ComE Operonprotein	for: TTATATCTATCGCACAGGCGGAGAG
		rev: GGTATACACCAGGCTTCTGAACAGC
<i>fdhD</i>	Formiat Dehydrogenase	for: AACTCGACCGTTAAGAAGCGTACAG
		rev: GGTATACACCAGGCTTCTGAACAGC
<i>gerBB</i>	Sporenkeimungsprotein B2	for: AGCTGTACTGGTTGGTACTGCGATT
		rev: GAGACAGTCCTAATGAAGCAGCGTA
<i>mmsA</i>	Methylmalonatsemialde- hyd Dehydrogenase	for: GCCGCTTTTAGCAGCAAGATTAGT
		rev: CAGCTACTGGCTGTGAACCTACAAA
<i>pfyP</i>	Blaulichtrezeptor	for: GTGCCAATAGAGGAAGGAATCTCTG
		rev: CTTGTGTTACGCGAGGTGAAGCTAT
<i>ymxH</i>	Sporulationsprotein	for: GAGCAGAGCGTTTGGGTATTTTAGG
		rev: CACTCGCACTTCATCTCCTTGTTT
<i>yoaK</i>	Membranprotein	for: TTATATCTATCGCACAGGCGGAGAG
		rev: GGTATACACCAGGCTTCTGAACAGC
<i>yoeB</i>	Membranprotein	for: CTTTACGTACAAAGTGGCGGTAAGG
		rev: GACTAGATGCTGTACTTGGCGCATT
GOI		
<i>xylA</i>	Xylose-Isomerase	for: CCCTCTTTTAGGCTGGGATACAGAT
		rev: CGTCTTGTTCAAAAGAACCTCTGC
<i>xylB</i>	Xylulose-Kinase	for: GTAGTATTGGGACATCGGGTGTGT
		rev: GCAGATAGCGTAACACCCATCGTAT
<i>xylR</i>	Xylose-Repressor	for: GCCTTGTAGATCGTCATCAGCAAA
		rev: CGTATGCTCCTGCATTAGCTTCAT
<i>xylT</i>	Xylose-Transporter	for: TTATATCTATCGCACAGGCGGAGAG
		rev: GGTATACACCAGGCTTCTGAACAGC
<i>gfp</i>	Grün fluoreszierendes Protein	for: TCTCGGACACAACTCGAGTACAAC
		rev: CTGCTAGTTGAACGGATCCATCTTC

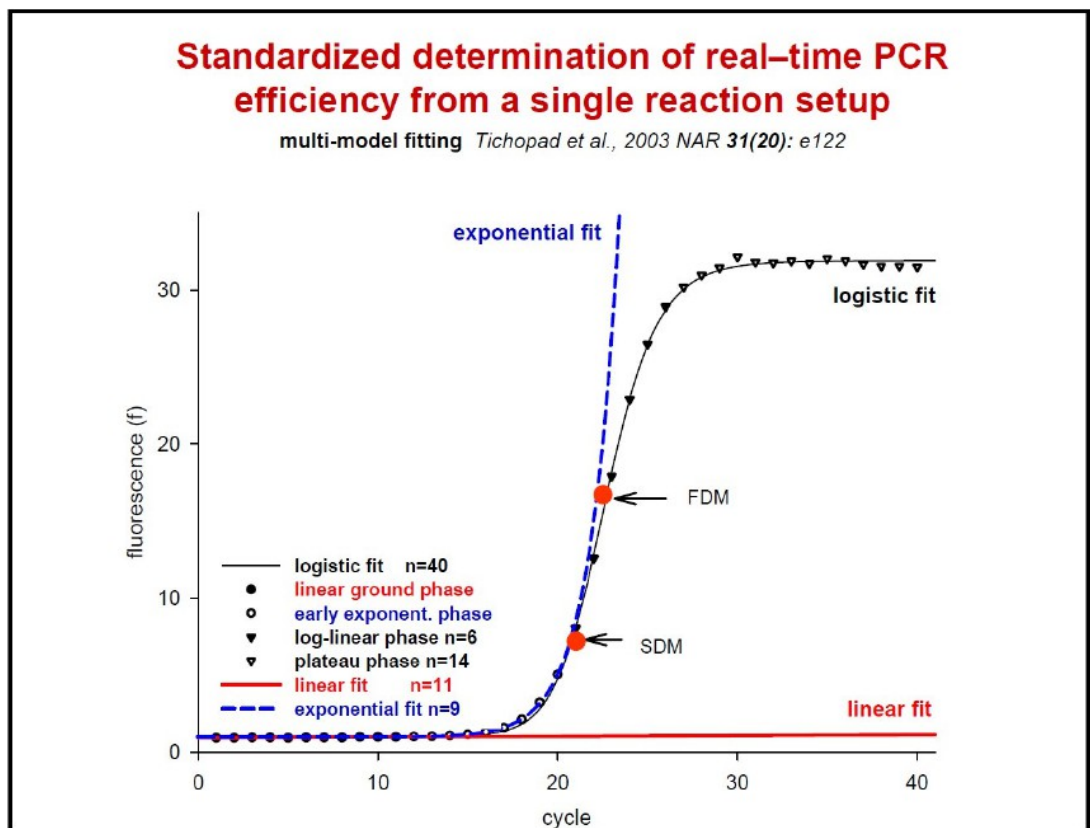


Abbildung 10: Vier-Parameter Modell von Tichopad *et al.*, 2003. Der relevante Bereich für die Effizienzberechnung ist die exponentielle Phase. Dafür werden Ausreißer gesucht, die nicht mehr dem linearen Fit angehören. Der ermittelte C_q -Wert entspricht dem Maximum der zweiten Ableitung (SDM). Bis zu diesem Punkt findet eine Verdoppelung des Amplikons statt, sofern die Effizienz bei 100 % liegt.

Die Plasmid Copy Number wurde nach Škulj *et al.*, 2008 bestimmt. Entweder wurden aufsteigende Mengen an grünen bzw. weißen Zellen eingesetzt (zwischen 100 bis 10.000) oder je Subpopulation eine definierte Konzentration an Gesamt-DNA. Die Zellzahlbestimmung erfolgte durch das Sortieren der Zellen oder anhand photometrischer Bestimmungen, da die OD auf ausgezählte Zellen mit Hilfe des Mikroskops kalibriert wurde. Dazu wurden die Zellen in 20 μ l H_2O_{deion} gelöst und 10 min bei 95° C inkubiert. Im Anschluss erfolgte ein 15-minütiger RNase-Verdau bei 37° C. Um inhibitorische Effekte auszuschließen, wurde eine Verdünnung von 1 : 100 durchgeführt und 1 μ l für die Real-Time PCR eingesetzt.

Die PCN wurde nach folgender Formel berechnet: $E_{\text{Chromosom}}^{C_q \text{ Chromosom}} / E_{\text{Plasmid}}^{C_q \text{ Plasmid}}$.

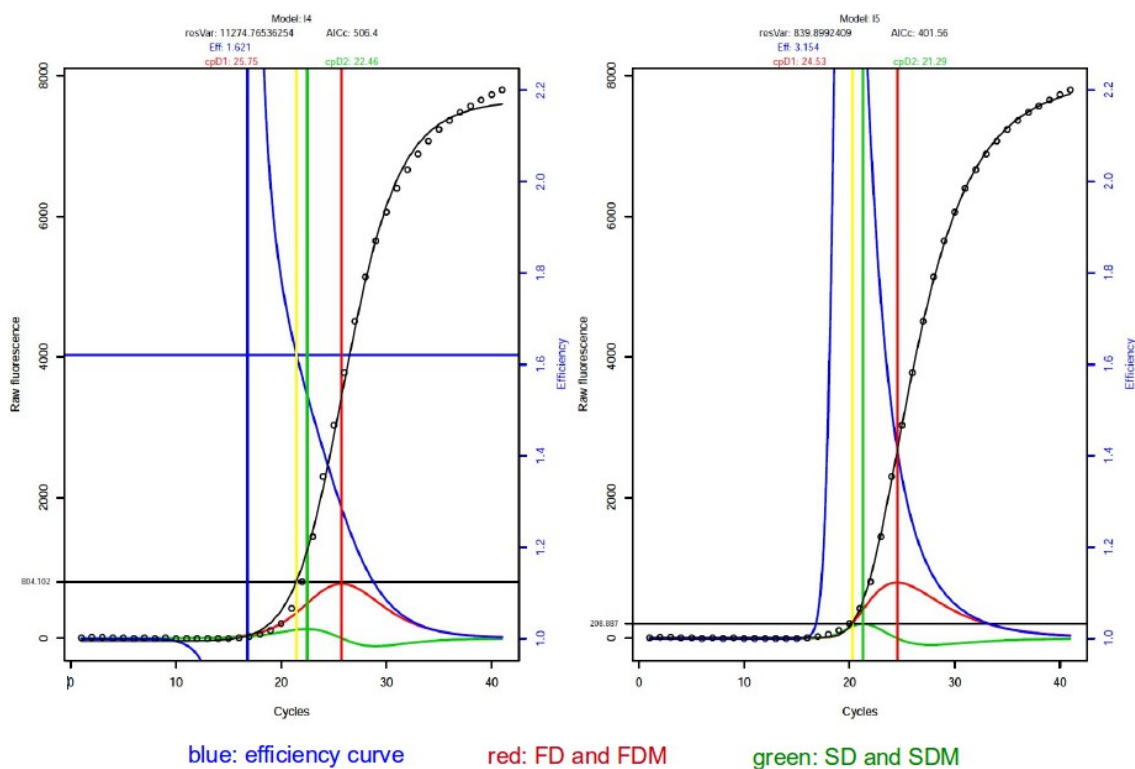


Abbildung 11: Vergleich der beiden log-logistischen Modelle von Ritz und Spiess, 2008. Das linke Diagramm zeigt das vier-Parameter Modell nach Tichopad *et al.*, 2003 und das rechte Diagramm das Modell mit fünf Parametern von Spiess *et al.*, 2008. Die blaue Kurve beschreibt die Effizienzberechnung, die rote die erste Ableitung (FD) und dessen Maximum (FDM), sowie die blaue die zweite Ableitung (SD) und das entsprechende Maximum (SDM).

2.7.11 Fluoreszenzmikroskopie und konfokale Laser-Scanning-Mikroskopie

Die in den letzten Jahren generierten Durchschnittsdaten einer Population sind nicht repräsentativ für eine einzelne Zelle (Verhalten, Status oder Phänotyp), so dass Techniken entwickelt wurden die nicht nur Wachstum, Teilung und Differenzierung einer einzelnen Zelle verfolgen, sondern auch deren Entwicklung und Abstammung analysieren (De Jong *et al.*, 2011). Eine solche Technik ist die quantitative Zeitraffer-Fluoreszenz-Mikroskopie. Mit der Hilfe von Fluoreszenzproteinen lassen sich Genexpressionen, Promotoraktivitäten oder Proteinlokalisationen verfolgen (Young *et al.*, 2012), so dass phänotypische Unterschiede einer genetisch identischen Population, epigenetische Stadien und biochemische Interaktionen sowie Zell- und Polalterung einer einzelnen Zelllinien analysiert werden können (Locke *et al.*, 2009).

Die verwendete Agarose-Pad Methode mittels Agaroseblock hat den Vorteil, genügend Medium für die Zellen über einen längeren Zeitraum bereitzustellen (Young *et al.*, 2012). Darüber hinaus ist aufgrund seiner geringen Größe im Verhältnis zum Objektträger (μ -Dish, 35 mm) für ausreichend Luftaustausch gesorgt (Young *et al.*, 2012).

Die Segmentierung (räumlichen Abgrenzung der Zelle vom Hintergrund) bezüglich der Zellteilung sollte mit Hilfe der Zeitraffer-Mikroskopie verfolgt und untersucht werden. Dies wurde mit einem reversen Fluoreszenzmikroskop (Axiovert 200M) und einer AxioCam CCD-Kamera von Zeiss sowie einem beheizbaren Mikroskop-Inkubator (37 °C) von ibidi verrichtet. Für die Lebendzellmikroskopie wurden nicht induzierte Zellen auf einen Agaroseblock aufgetragen (Young *et al.*, 2012), der aus A5-Medium zuzüglich 1,5 % Agarose sowie 10 μ g/ml Tetracyclin und 0,5 % Xylose hergestellt wurde. 1 μ l der Zellen, die bis zur $OD_{578} = 0,2$ gewachsen sind, wurden auf das Pad aufgetragen und mit der Kultur nach unten mittig auf ein μ -Dish (35mm Durchmesser) von ibidi aufgelegt. Mittels automatisierter Bildaufnahme wurde alle 5 min eine Zeitreihe aufgenommen. Dazu wurden die Kanäle Hellfeld, GFP und mCherry verwendet, deren Intensität zunächst in verschiedenen Graustufen aufgenommen wurde, um danach das Resultat in definierten Falschfarben veranschaulichen zu können. Je weiter dabei die Fluoreszenzspektrum auseinander liegen, desto besser ist der Unterschied der Fluoreszenzkanäle.

Die Bildanalysen wurde unter Verwendung von AxioVision 4.8 (Zeiss) und der TLM-Tracker-Software (Klein *et al.*, 2012) vorgenommen.

Bei der konfokalen Laser-Scanning-Mikroskopie werden die Zellen mit Hilfe eines Lasers über ein bewegliches Spiegelsystem abgetastet. Durch verschiedene Focusebenen können dreidimensionale Bilder aufgenommen werden. Diese Methode wurde mit induziertem pKMMBm2 bei der Firma Leica in Mannheim mit einem TCS SP8-Mikroskop durchgeführt. Zellen wurden auf einem Objektträger hitzefixiert und mit Mounting Medium von ibidi überzogen, um ein Ausbleichen der Fluoreszenz zu verzögern.

2.7.12 Durchflusszytometrie und FACS

Die Durchflusszytometrie ermöglicht die Quantifizierung von Zellen in einer Suspension aufgrund ihrer Form, Granularität und ihrer Färbung. Die Bakterienlösung, die zur Messung über eine dünne Kapillare durch eine Flusszelle (Flow Cell) gesogen wird, wird mittels Laser angeregt. Mit Hilfe der Kapillare passieren die Zellen als Einzelzellen die Messkammer. Dabei wird das Licht in verschiedene Richtungen gestreut. Das Vorwärtstreulicht (FCS = Forward Scatter) beschreibt die Beugung des ursprünglichen Strahls im flachen Winkel und ist proportional zur Zellgröße. Im Gegensatz dazu beschreibt das Seitwärtstreulicht (SSC = Sidewards Scatter) die Brechung und Reflexion im 90° Winkel zum Ursprungsstrahl. Somit verursachen kleine Zellen ein kleines Vorwärtstreulicht und große Zellen ein hohes Vorwärtstreulicht. Zellen mit geringer Granularität bewirken ein niedriges Seitwärtstreulicht, solche mit hoher Granularität ein großes. Darüber hinaus können durch den entsprechenden Laserstrahl Fluorochrome angeregt werden, so dass die Fluoreszenz jeder einzelnen Zelle bestimmt werden kann.

Auf diese Weise können innerhalb weniger Sekunden bis zu 100.000 Zellen gemessen werden. Daraus resultiert ein Streulicht-Dot-Plot, in dem jeder Punkt einer gemessenen Zelle oder einem gemessenen Ereignis entspricht.

Mit Hilfe des Durchflusszytometers Cube 8 von PARTEC wurden verschiedene Plasmide im *Bacillus megaterium* Wildtypstamm (DSM 319) analysiert. Die Bakterienkultivierung wurde wie unter Kap. 2.6.2 beschrieben vorgenommen. Die Probeentnahme erfolgte vor Induktion, sowie 1, 2, 3, 4 und 5 Std nach Zugabe von Xylose und endete mit einer letzten Entnahme der über Nacht gewachsenen Kultur. Mit Hilfe des Photometers wurde die Konzentration auf 10^6 Zellen/ml eingestellt. Um die Bakterienketten aufzubrechen, wurden die entnommenen Proben für 5 sek bei 70 % Power in einem 50 ml Falconröhrchen im Eisbad sonifiziert. Es erfolgte ein zweimaliger Waschschrift mit 1 x PBS bei 2100 rpm.

Zum Überprüfen der Vitalität von Zellen, wurde von einigen Proben eine Färbung mittels PI vorgenommen. Ist die Zelle tot, kann dieser Farbstoff durch die perforierte Zellmembran in die Zelle eindringen. Dazu wurde vor der Sonifikation und dem darauffolgenden Waschschrift die PI-Stocklösung (1 mg/ml in 1 x PBS) 1:500 in 1 x

PBS verdünnt, das Pellet in 1 ml Färbelösung resuspendiert und für 15 min bei RT im Dunkeln inkubiert.

Die Anregung der Proben fand im Blaulichtbereich bei 488 nm für eGFP statt und im Gelblichtbereich bei 561 nm für mCherry. Von jeder Probe wurden 100.000 Ereignisse detektiert und die Daten mit Hilfe der Programmiersprache R ausgewertet. Bioconductor bietet hierbei die R-Pakete flowCore und flowViz an (Hahne *et al.*, 2009; Sarkar *et al.*, 2008).

Eine besondere Form der Durchflusszytometrie ist das Fluorescence Activated Cell Sorting (FACS). Hierbei werden Zellen je nach Markierung in verschiedene Reaktionsgefäße sortiert.

Die Zellsortierung wurde mit Geräten der Firma Becton-Dickinson vorgenommen, einem FACSARIA-II und einem FACSARIA-II SORP. Dies geschah am Helmholtz-Zentrum für Infektionsforschung (HZI) mit freundlicher Unterstützung von Maria Höxter und Dr. Lothar Gröbe. Die Kultivierung der Bakterien wurde wie unter Kap. 2.6.2 beschrieben vorgenommen. 2 Std nach Induktion wurde die Kultur geerntet, mit 1 x PBS gewaschen und auf eine OD von 1,0 eingestellt.

2.7.13 FISH

Bei der Fluoreszenz-*in-situ*-Hybridisierung (FISH) können mit Hilfe einer fluoreszenzmarkierten Sonde Nukleinsäuren am Ursprungsort (*in situ*) mit Hilfe eines Fluoreszenzmikroskopes nachgewiesen werden. Die dabei verwendeten Sonden beinhalten komplementäre Nukleinsäuren im Bereich der Zelle, des Zellkerns, der Chromosomen oder der Gewebe die man nachweisen möchte. Die einzelsträngige Sonde hybridisiert dabei an die einzelsträngige Zielnukleinsäure. Um doppelsträngige Ziel-DNA zu denaturieren, kommen Reagenzien wie Formamid zum Einsatz. Besondere Waschschrte, die sich in ihrer Temperatur und Salzkonzentration unterscheiden, entfernen überschüssiges Sondenmaterial und unspezifische Bindungen.

Das benutzte FISH-Protokoll basiert auf der Methode nach Jensen und Shapiro, 1999 und wurde von Pogliano *et al.*, 2001 optimiert. Zur Herstellung der Sonde (Dernburg and Sedat, 1998) wurde Plasmid-DNA von pRBBm99 oder pBC16 (Bernhard *et al.*,

1978) unter Verwendung von Enzymen, deren Erkennungssequenz vier Basenpaare beträgt (*AluI*, *HaeIII*, *MseI*, *MspI*, *RsaI* und *Sau3AI*), verdaut. Dabei wurden kleine Fragmente im Bereich von 20 bp bis 280 bp generiert. Im Anschluss wurde eine Phenol-Chloroform-Extraktion durchgeführt und die DNA-Fragmente über eine terminale Transferase (New England BioLabs) mit Cy3-dCTP (GE Healthcare) markiert.

Die Bakterienzellen wurden wie unter Kapitel 2.6.2 (Kultivierung von Bakterien) und 2.6.3 (Induktion der plasmidtragenden Stämme) beschrieben kultiviert und 1 Std nach Induktion geerntet. Als Kontrollen dienten die nicht induzierten Stämme. Die Suspension wurde mit 0,87 % Paraformaldehyd und 1 M NaPO₄-Puffer (pH 7,4) für 10 min bei RT fixiert. Danach erfolgte ein dreimaliger Waschschrift mit 1 x PBS und ein Resuspendieren der Zellen in 50 µl GTE zuzüglich 10 µl Lysozym (2 mg/ml). 10 µl des Ansatzes wurde für 10 min auf einen Poly-L-Lysine beschichteten Objektträger von ibidi (Poly-L-Lysine Slide) gegeben, mit 1 x PBS gewaschen und für 3 min bei RT getrocknet. Um unspezifische DNA-Bindungen zu verhindern und das Hintergrundrauschen zu minimieren, wurde mit Heringssperma-DNA geblockt. Dafür wurde die Lösung für 3 min auf 100 °C erhitzt, sofort auf Eis gestellt und je 10 µl auf jede Probe gegeben. Der Objektträger wurde für 2 min bei 75 °C erhitzt und direkt für 5 min in ein eiskaltes EtOH-Bad (70 %) gegeben. Von dort wurde er direkt für weitere 5 min in ein 90 %iges EtOH-Bad transferiert und zum Schluss für 5 min zu eiskaltem absoluten EtOH. Der Slide wurde für 5 bis 10 min bei RT getrocknet. Es erfolgte eine Zugabe von 2 x SSC und 50 % Formamid. Nach 5 min wurde die überschüssige Lösung entfernt. In dieser Zeit konnte die Sonde vorbereitet werden. Die Lösung für die Sonde bestand aus 50 % Formamid, 20 x SSC sowie 0,5 µg/ml Heringssperma-DNA, Sonden-DNA und 4,375 % Dextransulfat. Im Anschluss wurde die Sonde für 10 min bei 75 °C erhitzt. Die Konzentration betrug dabei 200 ng/µl. Die Sonde wurde auf die Zellen aufgeteilt, für 2 min bei 94 °C erhitzt und über Nacht bei 42 °C hybridisiert. Nach 8 Std wurden die Zellen zunächst mit 2 x SSC und 50 % Formamid für 30 min bei 42 °C gewaschen. Es folgte ein weiterer Waschschrift (2 x SSC und 25 % Formamid) für 10 min bei RT. Anschließend wurde ein dreimaliges Waschen der Zellen (2 x SSC) für jeweils 10 min bei RT vorgenommen. Die beiden letzten Waschschriffe wurden mit 1 x PBS durchgeführt.

Für die DAPI-Färbung wurden die Zellen in 300 nM DAPI-Dihydrochlorid von

Invitrogen bei RT inkubiert, dreimal mit A5-Medium gewaschen und zu den Proben gegeben.

2.8 Software

Für die Auswertung der Daten wurde die freie Programmiersprache R (<http://www.r-project.org/>) und folgende frei erhältlichen Zusatzpakete verwendet:

Biobase	Bioconductor, Huber et al.
deSolve	cran.R-project, Soetaert, K., Petzoldt, T. and W.R., Setzer
diptest	cran.R-project, Maechler, M.
flowCore	Bioconductor, Ellis et al.
flowViz	Bioconductor, Ellis et al.
geNorm	Bioconductor, Vandesompele et al.
gplots	cran.R-project, Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Liaw, W.H.A., Lumley, T., Mächler, M., Magnusson, A., Möller, S., Schwartz, M. and Venables, B.
nlrwr	cran.R-project, Ritz, C.
plotrix	cran.R-project, Lemon, J., Bolker, B., Oom, S., Klein, E., Rowlingson, B., Wickham, H., Tyagi, A., Etteradossi, O., Grothendieck, G., Toews, M., Kane, J., Turner, R., Witthoft, C., Stander, J., Petzoldt, T., Duursma, R., Biancotto, E., Levy, O., Dutang, C., Solymos, P., Engelmann, R., Hecker, M., Steinbeck, F., Borchers, F., Singmann, H, Toal, T. and Ogle, D.
qpcR	cran.R-project, Spiess, A.N.
SlqPCR	Bioconductor, Kohl et al.

2.8.1 BioLector

Die Rohdaten des BioLectors wurden mittels der zugehörigen BioLection-Software von m2p-labs ausgewertet.

2.8.2 Microarray

Für die Arrayauswertung wurde das Bioconductorpaket limma (Linear Models for Microarray and RNA-Seq Data) verwendet. Die Auswertung beinhaltete mehrere Schritte. Als Erstes wurden die Microarray-Rohdaten des Array-Scanners geladen. Zu diesen Daten gehört das Array-Layout, das dem jeweiligen Genspot seine Position zuordnet. Darüber hinaus muss das experimentelle Design übertragen werden. Dieses beinhaltet die Anzahl der experimentellen Replikate und untersuchten Zustände, sowie die benutzte Fluoreszenzfarbstoffmarkierung (Cy3 bzw. Cy5). Im nächsten Schritt wurden die Expressionsverhältnisse berechnet und eine \log_2 Transformation der Rohdaten durchgeführt. Mit diesen Daten wurde dann eine Hintergrundkorrektur und Normalisierung zwischen den Replikaten vorgenommen. Hierbei wurde die "Loess" Regressionsmethode und Quantile-Normalisierung verwendet (Yang and Thorne, 2003). Im letzten Schritt konnten über ein lineares Modell Standardfehler und p-Werte berechnet werden. Als Endresultat enthielt man eine Tabelle mit allen Genen, sowie den berechneten und korrigierten Expressionsverhältnissen (FC bedeutet Fold Change). Zum Erstellen des Venn-Diagramms wurde aus dem CRAN-Archiv (Comprehensive R Archive Network) der Pakete - basierend auf R - auf das Package VennDiagram von Hambo Chen zurückgegriffen.

2.8.3 qRT-PCR

Die Rohdaten der qRT-PCR wurden über die Firmensoftware CFX Manager von BioRad exportiert und in ein CSV-Format (Comma Separated Values) umgewandelt. Bei diesem Format werden die Inhalte einer Zeile durch Koma oder Semikolon getrennt. Anschließend wurden die Daten mit dem qpcR-Paket (Ritz and Spiess, 2008) analysiert.

Zum Ermitteln der relativen Menge an cDNA ist eine Normalisierung der Zielgene mit Hilfe von Referenzgenen unter Einberechnung der Effizienz erforderlich. Als interner Standard dienen Gene, die grundlegende Zellfunktionen besitzen und auch bei unterschiedlichen Versuchsbedingungen nicht variieren. Dabei sollen keine signifikan-

ten Expressionsunterschiede zwischen der behandelten und der unbehandelten Probe aufgezeigt werden, was die Benutzung eines einzelnen Referenzgenes unmöglich macht (Tricarico *et al.*, 2001 und Thellin *et al.*, 1999). Programme wie NormFinder, BestKeeper und geNorm basieren auf verschiedenen Algorithmen und ermöglichen dadurch die Berechnung mehrerer Referenzgene. Der enorme Vorteil von geNorm liegt in der zusätzlichen Berechnung der optimalen Anzahl an Referenzgenen.

Alle ermittelten C_q -Werte wurden unter Einberechnung der Effizienz E nach folgender Formel in relative Expressionswerte (W_R) umgerechnet: $W_R = E^{(\min C_q \text{ Mittel} - C_q \text{ Mittel})}$, wobei im konkreten Beispiel $\min C_q \text{ Mittel}$ der niedrigste der gemittelten Werte innerhalb einer Kultur mit oder ohne Plasmid zu unterschiedlichen Bedingungen (vor und nach Induktion der verschiedene Stämme) bedeutet und $C_q \text{ Mittel}$ der gemittelte Wert innerhalb dieser verwendeten Kulturen meint. Für diese Berechnungen wird jeweils das arithmetrische Mittel benötigt. Dabei wird dem niedrigsten C_q -Wert automatisch die Zahl eins zugeordnet (Vandesompele *et al.*, 2002). Im Anschluss wurden die ermittelten Werte in das Programm importiert.

Eine geometrische Normalisierung wurde unter Benutzung der Software geNorm (Vandesompele *et al.*, 2002) validiert. Die Datenanalyse erfolgte mit dem Bioconductor R-Package SLqPCR (Kohl, 2007). Die Stabilität der Kandidatengene wird paarweise über den Stabilitätsfaktor M berechnet. Je kleiner dieser Wert ist, desto stabiler ist die durchschnittliche paarweise Variation eines Referenzgens im Vergleich zu den anderen analysierten potenziellen Referenzgenen. Das Gen mit dem jeweils höchsten M -Wert wird dabei aus den folgenden Berechnungen ausgeschlossen, bis nur noch zwei Referenzgene übrig bleiben. Daraus resultiert ein stabilitätsabhängiges Ranking. Darüber hinaus wird der V_{NF} -Wert (Varianz des Normalisierungsfaktors) berechnet, der die optimale Referenzgenanzahl zur Normalisierung anzeigt. Dabei gilt, je niedriger der V_{NF} -Wert, umso geringer der Einfluss auf die Normalisierung, den ein zusätzliches Referenzgen hat. Der empfohlene Wert konnte in allen Berechnungen erreicht werden und lag bei $< 0,15$. Eine relative Quantifizierung mit den entsprechenden Normalisierungsfaktoren konnte nun durchgeführt werden.

2.8.4 Time-Lapse Mikroskopie

Die Bearbeitung und der Export der Rohdaten des Mikroskopes im ZVI-Format erfolgte mit der Software AxioVision von Zeiss. Die exportierten Einzelbilder wurden mit TLM (Time-Lapse-Movie)-Tracker (Klein *et al.*, 2012) ausgewertet. Das auf der MATLAB-Plattform basierende Programm bietet über Bildanalyseverfahren die Möglichkeit, die Auswertung der Zeitraffer-Mikroskopiedaten für Einzelzellanalysen durchzuführen. Es ermöglicht die flexible und benutzerfreundliche Segmentierung (räumlichen Abgrenzung der Zelle von ihrem Hintergrund), Verfolgung (Tracking) und Analyse von mikrobiellen Zellen. Durch die Verwendung von TLM-Tracker ist es möglich, räumliche und zeitliche Ereignisse wie Wachstumsraten und Fluoreszenzintensitäten einer Zellteilung zu verfolgen und zu messen. Die gewonnenen Daten können zur Generierung und Analyse von Zelllinienbäume verwendet werden. Die bimodale Verteilung der Zelllinien wurden unter Verwendung des Diptest R Pakets von Martin Maechler analysiert.

Die Software TLM-Tracker und ihre dazugehörige Bedienungsanleitung ist unter <http://www.tlmtracker.tu-bs.de> abgelegt.

2.8.5 Durchflussszytometrie

Die Durchführung der Messung erfolgte durch das Programm CyView. Die Rohdaten wurden im FCS-Format exportiert (Bray *et al.*, 2012). Weil das benutzte Durchflussszytometer ein Datenformat generiert, das mit der Benutzung in R nicht kompatibel ist, wurde zur Problemlösung ein Skript entworfen. Für die statistische Auswertung wurde das R Package FlowCore und FlowViz genutzt. Die Auswertung beinhaltet eine log-Transformation der Daten, die Visualisierung als Dichteplot der Daten, die Definition von Zellgruppen in Regionen (Gates), sowie die Quantifizierung von Zellen.

3. Ergebnis und Diskussion

3.1 Makroskopische und mikroskopische Untersuchungen

Die beschriebene Kulturheterogenität von induzierten, eGFP-produzierenden Zellen konnte auch bei Kolonien beobachtet werden. Der Stamm, der das Produktionsplasmid pSSBm85 trug, wurde auf xylosehaltigem A5-Medium ausplattiert. Daraufhin ergab sich auf der Platte eine Segregation zweier coexistierender Subpopulationen. Auf Festmedium wurden in etwa 70 % kleine, grüne, hochproduktive sowie 30 % große, weiße, schwach produzierende Kolonien gefunden (Abb. 12A).

Die mikroskopische Untersuchung ergab, dass weiße Kolonien noch eine geringe Menge an eGFP erzeugten, jedoch war die Fluoreszenz im Vergleich einige hundertmal niedriger als in Zellen aus grünen Kolonien. Diese Assoziation geht aus Abbildung 12B hervor.

Phänotypisch betrachtet wiesen alle verwendeten plasmidtragenden Stämme eine Zellheterogenität auf. Auf diese Beobachtung soll später genauer eingegangen werden.

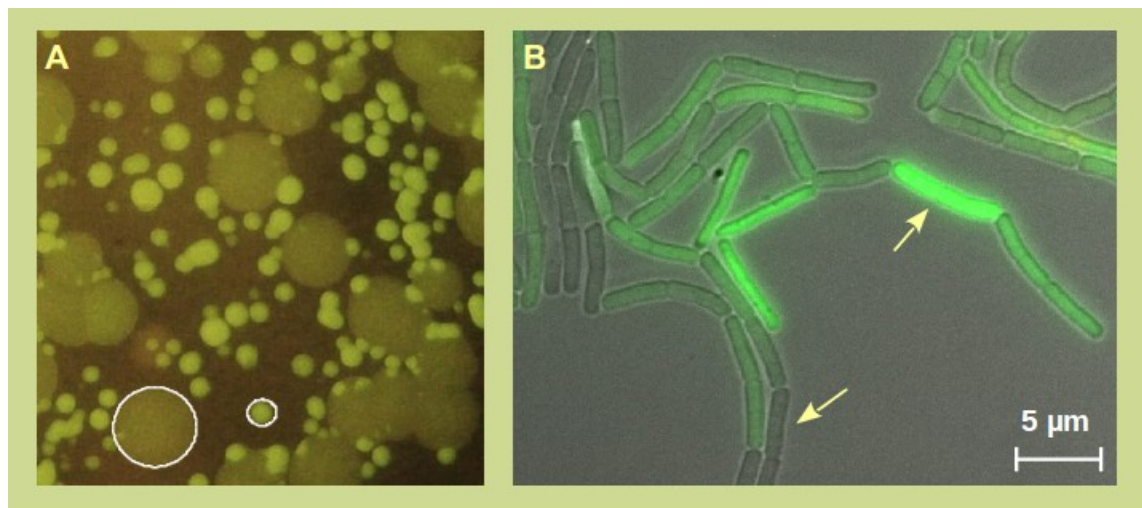


Abbildung 12: Kolonieheterogenität in *B. megaterium* des eGFP produzierenden Stamms pSSBm85.
A) Unter dem Blaulicht-Transilluminator weist der große Kreis auf ein Beispiel für eine weiße Kolonie hin und der kleine Kreis auf eine coexistierende, grüne, stark eGFP-produzierende Kolonie. B) Innerhalb der fluoreszenzmikroskopischen Aufnahme einer wachsenden Mikrokolonie deutet der linke Pfeil auf eine nicht- bzw. wenigproduzierende Zelle und der rechte auf eine grünfarbene, stark eGFP-produzierende Zelle.

Um ausschließen zu können, dass der beobachtete Effekt auf der Folge eines Plasmidverlusts während der Kultivierung beruhte, wurden beide Zelltypen nochmals auf antibiotikumhaltiges Medium transferiert. Ein erneutes Ausplattieren beider Populationen bedeutete, dass weiße Kolonien fast ausschließlich (98 % - 99 %) weiß blieben, wohingegen grüne Kolonien abermals im selben Verhältnis, wie oben beschrieben, segregierten. Das gleiche Resultat konnte auch bei den beiden darauf folgenden Generationen bestätigt werden. Tendenziell blieben weiße Kolonien weiß. Ein sporadisches Auftauchen von grünen Kolonien (1 % - 2 %) konnte in den Folgegenerationen nicht analysiert werden, da diese Kolonien extrem winzig waren und eine Weiterkultivierung dadurch nicht möglich war. Sortierte Zellen mittels FACS resultierten in einer grünen und einer weißen Subpopulation. Kultivierte man diese beiden Fraktionen weiter, konnte wiederum ein ähnliches Muster registriert werden. Die Zahlenverhältnisse blieben bestehen. Bei den weißen, wenigproduzierenden Kolonien handelte es sich demzufolge noch um vitale Zellen was bedeutet, dass eine Subpopulation in einer Kultur aus klonalen, antibiotikaresistenten Zellen entstand. Zudem ließ sich aus diesen weißen Zellen das Plasmid isolieren, so dass eine genomische Integration ausgeschlossen werden konnte. Darüber hinaus konnte durch die Proliferation der weißen Zellen eine Persistenz ausgeschlossen werden. Dieses Ergebnis deutete auf ein bistabiles System hin.

Unter Verwendung unterschiedlicher Xylosekonzentrationen (0,1 % / 0,5 % / 1%) konnte kein Effekt bezüglich der Segregationsverhältnisse beobachtet werden, allerdings nahm die Koloniegröße mit erhöhter Dosis an Xylose ab, so dass im weiteren Verlauf 0,5 % als optimale Menge gewählt wurde. Die Resultate deuteten darauf hin, dass die Koloniegröße mit der Produktionsrate negativ korrelierte.

Bistabile Systeme sind in der Lage, Hysterese auszubilden. Dadurch kann ein physiologischer Zustand auch nach dem Wegfall einer Stimulierung noch andauern. Die Entscheidung der Zelle welcher Zustand gewählt wird, hängt hierbei von der Vorgeschichte ab (Solopova *et al.*, 2013). Das Produktionsplasmid pSSBm85 blieb in einem Plasmid Loss Versuch nach Bron und Luxen, 1985 auch ohne Selektionsdruck über mehrere Stunden stabil in *B. megaterium* bestehen. Dieser Versuch diente als Grundlage für weitere Analysen bezüglich des Plasmids pSSBm85. Eine schematische Darstellung der Vorgehensweise ist Abbildung 13 zu entnehmen.

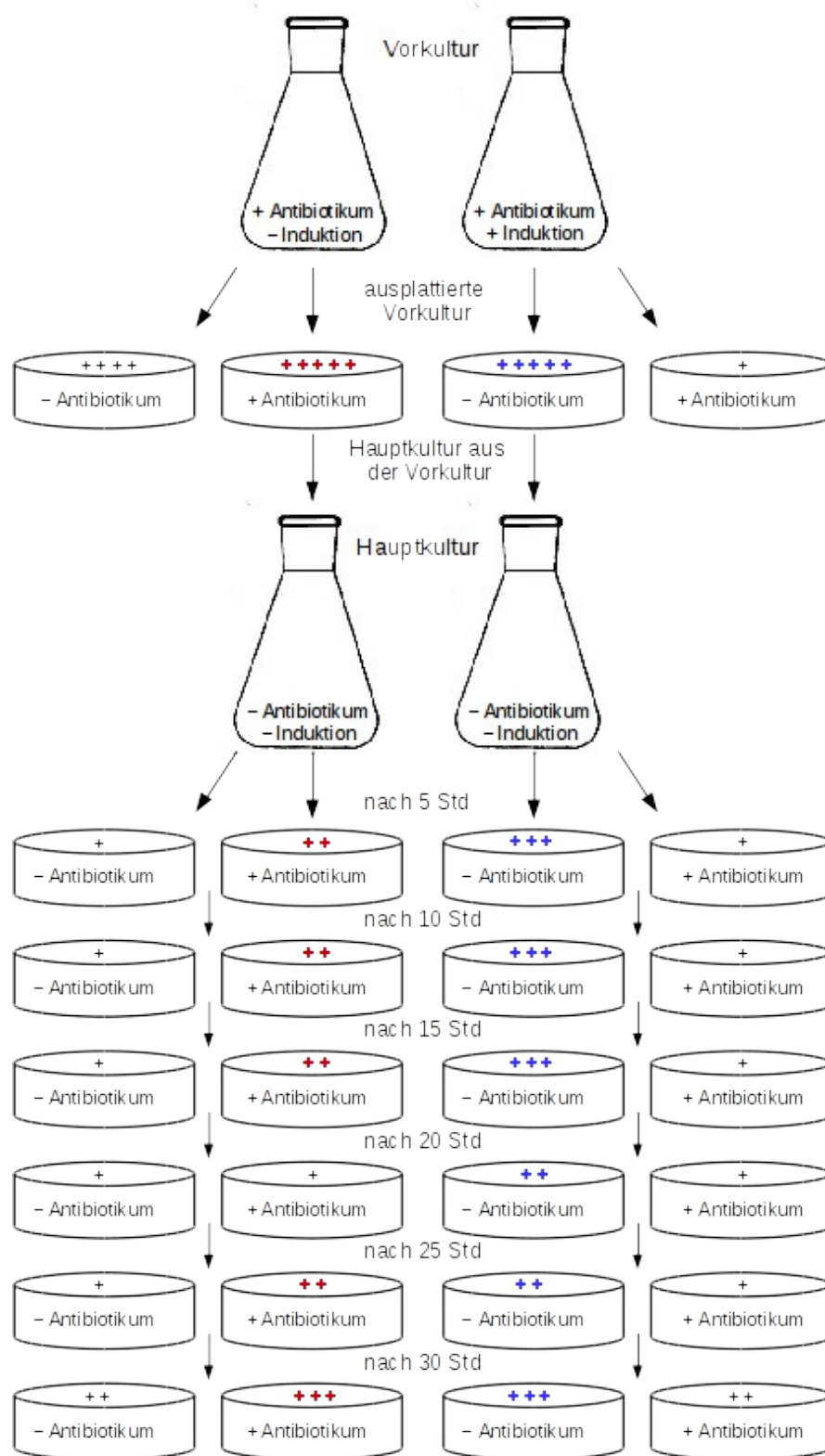


Abbildung 13: Versuchsschema zur Analyse der Heterogenität in *B. megaterium* mit dem Produktionsplasmid pSSBm85. Ausgehend von zwei Vorkulturen, die jeweils Tetracyclin enthielten wurden Kolonien ausplattiert. Eine der beiden Kulturen war bereits induziert (+ rechts im Bild), wohingegen die andere nicht induziert war (- links im Bild). Beide Bakteriensuspensionen wurden sowohl auf antibio-

tikumhaltigen, als auch auf antibiotikumfreien Medium ausplattiert. Die Hauptkulturen, die weder Tetracyclin, noch Xylose enthielten, unterschieden sich demnach in ihrer Vorgeschichte bezüglich der Induktion. Alle 5 Std wurden den Kulturen Aliquots entnommen und diese zum einen wie oben beschrieben ausplattiert und zum anderen in frisches Medium überführt und für weitere 5 Std inokuliert, bis sich der Vorgang wiederholte. Die Anzahl der plus-Zeichen ist proportional zur gewachsenen Kolonienmenge. Je mehr Kolonien vorhanden waren, desto größer die Anzahl der Zeichen. Prinzipiell bedeutet rot eine Vorgeschichte ohne Xylose, während blau aus einer ehemals induzierten Kultur hervorgeht. Die entsprechende Farbe befindet sich stets auf der Platte, bei der die meisten Kolonien gezählt wurden.

Um die Vorgeschichte einer Kultur zu betrachten, diente als Vorkultur sowohl eine induktorfreie, als auch eine bereits induzierte Bakteriensuspension. Beide Vorkulturen wuchsen in tetracyclinhaltigem Medium. Der Hauptkultur hingegen wurde kein Antibiotikum zugesetzt. Eine Zugabe von Xylose fand nicht statt. Im weiteren Verlauf wurde alle 5 Std die Population neu überimpft und Aliquots sowohl auf Medium mit, als auch ohne Tetracyclin inokuliert. Für diesen Versuch sollen bewusst keine Zahlen genannt werden, da die Anzahl der Kolonien beider Kulturen zu sehr auseinanderdriftete. Bei der Vorkultur die nicht induziert wurde, konnte in der Hauptkultur stets ein besseres Wachstum bei gleicher Anfangs-OD und gleicher Wachstumszeit festgestellt werden, als bei der Versuchsreihe mit der induzierten Vorkultur. Handelte es sich bei der Hauptkultur um die, deren Vorkultur xylosefrei war (links im Bild), befanden sich durchweg (auch am Ende der Vorkultur) weniger Kolonien auf der Platte ohne Selektionsdruck, als auf der mit Tetracyclin. Erwarten würde man bei einem Plasmidverlust genau das Gegenteil. Dies beweist die Plasmidstabilität, da selbst die Vorkultur mit Tetracyclin das selbe Zahlenverhältnis aufzeigte, wie eben beschrieben. Betrachtet man deshalb in der Vorgeschichte eine ehemals induzierte Kultur (rechts im Bild), so fand man in der xylosefreien Hauptkultur deutlich mehr Kolonien ohne Tetracyclin (Faktor 4-fach bis 62-fach), als mit Antibiotikum. Da die Vorkultur Tetracyclin enthielt und das selbe Resultat lieferte, würde dies einem möglichen Plasmidverlust widersprechen. Im weiteren Versuchsverlauf stieg zunächst der Anteil der weißen Kolonien im antibiotikafreien Medium beständig von 20 % auf 75 % an und blieb bis zum Versuchsende bestehen. Bei den Aliquots, die hingegen auf Tetracyclin wuchsen, traten zunächst zu 99 % grüne Kolonien auf. Nach 25 Std kletterte die Anzahl der weißen Zellen erstmals auf 10 %, wobei die Gesamtzellzahl im ähnlichen Verhältnis anstieg was den Schluss zuläßt, dass das Auftreten weißer Zellen mit beschleunigtem Wachstum einhergeht.

3.2 FACS-Analysen von Subpopulationen, die viel und wenig eGFP produzieren

Um weitere Informationen über die molekularen Grundlagen der Kulturheterogenität zu generieren, wurden induzierte Zellen sortiert und mittels qRT-PCR die Induktion der Transkripte von potenziell am Expressionssystem beteiligten Schlüsselgenen untersucht. Der Scatterplot einer FACS-Analyse ist in Abbildung 14 dargestellt. Die bimodale Verteilung der Zellen wurde durch die zwei vorhandenen Zellfraktionen bestätigt, die sich erheblich in der Produktionsrate bezüglich eGFP unterschieden. Die Reinheit der sortierten Zellen wurde durch erneutes Sortieren (Resorting) der getrennten Population bestätigt.

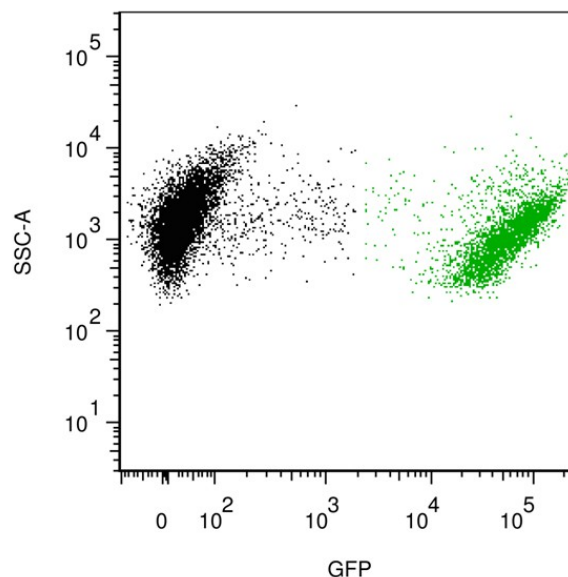


Abbildung 14: FACS-Analyse von eGFP-exprimierenden Zellen, die das Plasmid pSSBm85 enthielten. Die Zellsortierung einer Xylose-induzierten Kultur von *B. megaterium* ergab zwei Fraktionen, die geringproduktive (in Schwarz dargestellt) und leistungsstarke (abgebildet in Grün) Zellen widerspiegelten.

3.3 Analyse der Genexpression von *xylR*, *xylT* und *egfp* in den Subpopulationen mittels qRT-PCR

Die Expression der am Plasmidsystem beteiligten Gene der beiden Subpopulationen, die aus den sortierten Bakterienzellen resultierten, wurden mittels quantitativer Echtzeit-PCR analysiert. Dies betraf zunächst die beteiligten Gene *xylT*, *xylR* und *egfp*. Die RNA wurde wie in Kap. 2.7.2 beschrieben extrahiert.

3.3.1 Reinheitsbestimmung der RNA

Um die Zuverlässigkeit von qPCR-Daten zu steigern und fehlerhaften Ergebnissen vorzubeugen, die zu erheblichen Schwankungen der Genexpressionen führen würde, muss die Qualität der RNA mittels Kapillarelektrophorese überprüft werden (Fleige and Pfaffl, 2006). Die RIN-Werte wurden mittels Bioanalyzer ermittelt. Sie lagen zwischen 8,2 und 10,0 und entsprachen den Anforderungen der MIQE-Richtlinien. Die Qualitätsergebnisse der oben beschriebenen Analyse von extrahierter RNA aus dem Stamm DSM 319, der das Plasmid pSSBm85 trägt ist in Abbildung 15 abgebildet.

3.3.2 Ermitteln der Referenzgene

Das Programm geNorm diene zum Bestimmen der Referenzgene. In diesem Algorithmus wird zunächst eine stabilitätsabhängige Rangliste und die durchschnittliche Expressionsstabilität M erstellt. Der Vorteil dieser Berechnung liegt in der zusätzlichen Ermittlung der optimalen Anzahl der zu benötigenden Referenzgene. Dieser sogenannte V_{NF} -Wert gibt an, welchen Einfluss das in der Rangfolge nächstfolgende Referenzgen auf die Normalisierung hat. Je niedriger dieser Wert ist, desto geringer ist sein Einfluss. Empfohlen wird ein Schwellenwert von $V_{NF} \leq 0,15$ (Vandescompele *et al.*, 2002).

Für die nachstehende Normalisierung der RNA von pSSBm85 wurden folgende Gene in geNorm getestet: *rpoB*, *gyrB*, *heli1*, *heli2*, *gatB*, *gap2*, und *gap1*. Zwei Referenzgene würden laut Abbildung 16 ausreichen, wobei aber mindestens drei empfohlen werden (Vandesompele *et al.*, 2002). Das dritte und vierte Gen variierte mit den biologischen Replikaten, so dass mit vier Genen referenziert wurde (*rpoB*, *gyrB*, *heli1* und *heli2*). Eine genauere Beschreibung dieser Gene ist Tab. 10 in Kap. 2.7.10 zu entnehmen.

Unter Benutzung erweiterter Bedingungen, wie z. B. die verschiedenen Zeitpunkte nach der Induktion sowie der Verwendung anderer Plasmide in DSM 319 bzw. des leeren Plasmidstamms, waren diese Referenzgene nicht mehr ausreichend, so dass neue Referenzgene gefunden werden mussten. Als Hilfe für die Vorauswahl potenzieller Gene diene, angelehnt an die Idee von RefGenes - ein Programm von Genevestigator - (Hruz *et al.*, 2011), der in dieser Arbeit durchgeführte Microarray.

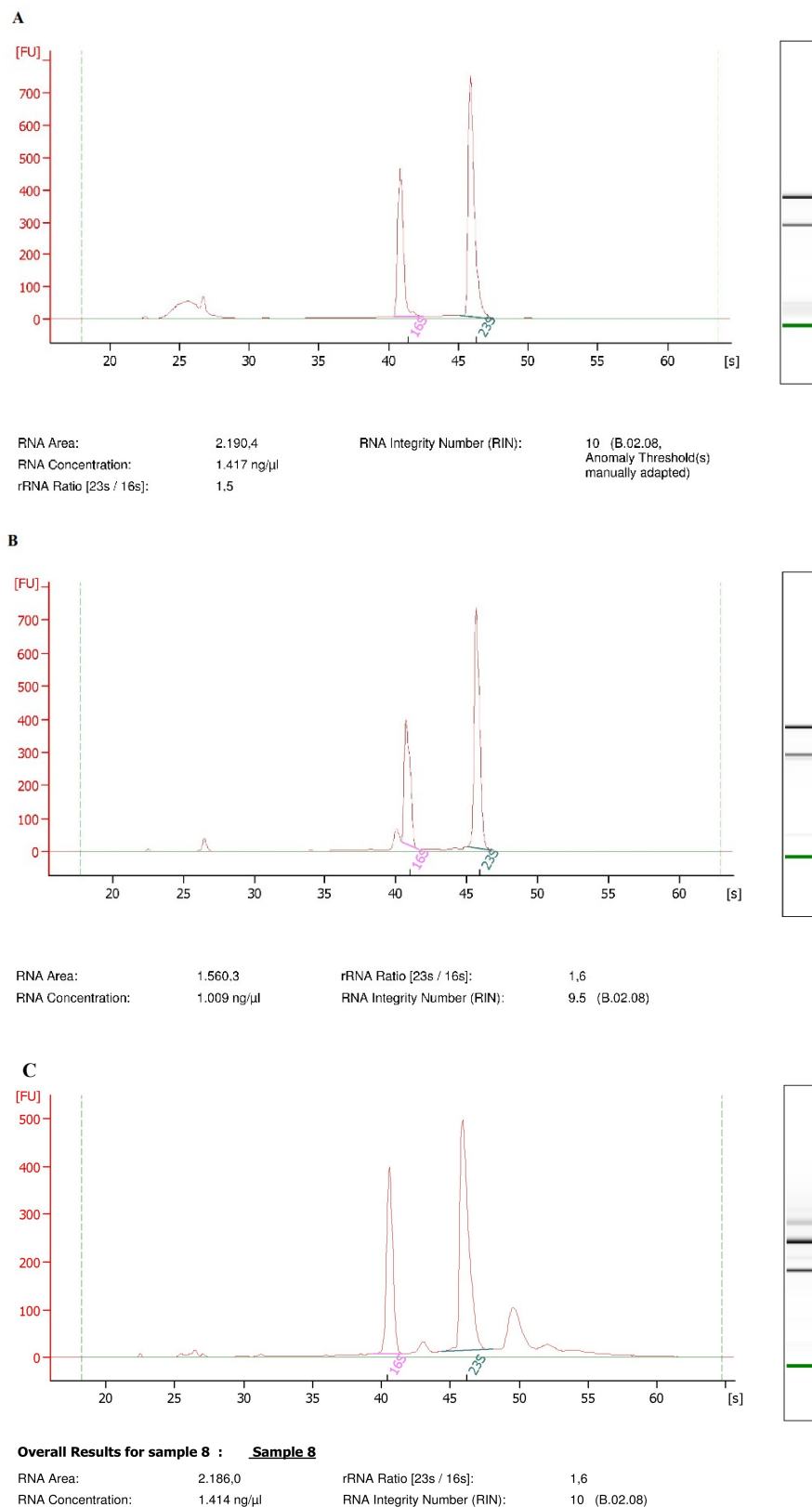


Abbildung 15: Qualität der RNA aus *B. megaterium* mit dem Plasmid pSSBm85 mittels Bio-analyzer. Elektropherogramme mit RIN-Werten von Proben A) vor Induktion B) FACS-sortierten niedrig-produzierenden Zellen und C) FACS-sortierten stark eGFP-produzierenden Zellen.

Die Berechnung bezüglich der Stabilität M und der Anzahl potenzieller Referenzgene ist Abbildung 17 zu entnehmen. Je höher dabei der Wert des Stabilitätsfaktors M ausfällt, desto instabiler ist die Genexpression unter verschiedenen Bedingungen. Das jeweils instabilste Gen fällt bei der Berechnung zuerst raus, so dass eine stabilitätsabhängige Auflistung erfolgt. Gleichzeitig wird die paarweise Variation (V_{NF} -Wert) eines zusätzlichen Gens im Vergleich zu den bereits analysierten ermittelt. Die Bewertung beinhaltet alle vorhanden plasmidhaltigen Stämme, den Wildtypstamm und deren biologischen Replikate, die zu unterschiedlichen Zeitpunkten nach der Induktion geerntet wurden. Die daraus resultierende Reihenfolge lautet: *aepY*, *comEA*, *ymxH*, *gerBB*, *yoeB*, *mmsA*, *pfyP*, *yoaK*, *fdhD* und *argR*. Gemäß Abbildung 18 würden laut der Varianz des Normalisierungsfaktors (V_{NF}) zwei Referenzgene zur Normalisierung ausreichen, wobei mindestens drei empfohlen werden (Vandesompele *et al.*, 2002). Demnach wurden die Gene *aepY*, *comEA* und *ymxH* ausgewählt. Eine genauere Beschreibung der Gene ist Tab. 10 in Kap. 2.7.10 zu entnehmen.

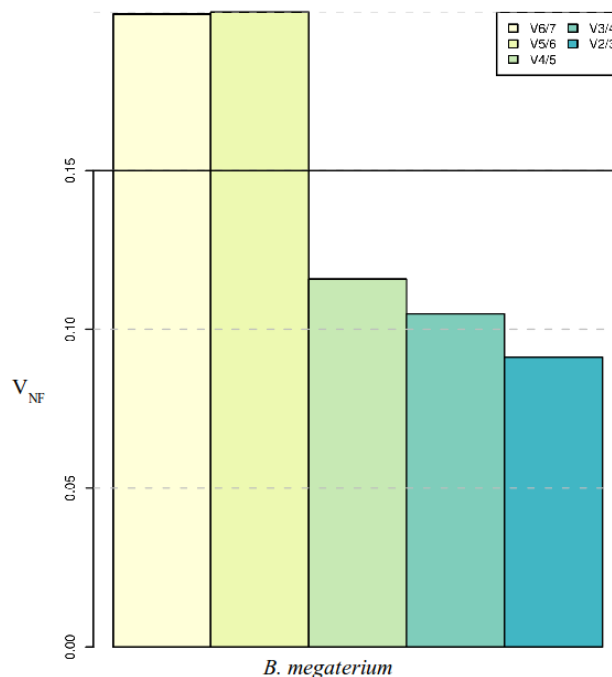


Abbildung 16: Anzahl der nach Stabilität geordneten Referenzgene des induzierten Stammes pSSBm85 mittels geNorm. Der V_{NF} -Wert gibt die Variation an, den ein zusätzliches Referenzgen auf die Normalisierung ausübt. $V2/3$ gibt die Variation zweier Gene in Bezug auf das nächst kommende dritte Gen in der Rangliste an. Zwei Gene würden für die Normalisierung ausreichen, wobei mindestens drei empfohlen werden. Der waagerechte Strich zeigt den von geNorm empfohlene V_{NF} -Wert von $\leq 0,15$ an.

```

Step 1 :
gene expression stability values M:
comEA ymxH aepY pfyP mmsA yoaK gerBB yoeB
0.6787731 0.6855744 0.7114827 0.7583553 0.7752470 0.7983887 0.8958062
0.9167455
fdhD argR
1.0775728 1.3642683
average expression stability M: 0.8662214
gene with lowest stability (largest M value): argR
Pairwise variation, ( 9 / 10 ): 0.1311535
#####
Step 2 :
gene expression stability values M:
comEA ymxH aepY mmsA pfyP gerBB yoaK yoeB
0.5817365 0.5860574 0.6076366 0.6822482 0.7416740 0.7807747 0.7921101
0.7953637
fdhD
1.1077859
average expression stability M: 0.7417097
gene with lowest stability (largest M value): fdhD
Pairwise variation, ( 8 / 9 ): 0.1167093
#####
Step 3 :
gene expression stability values M:
comEA aepY ymxH mmsA gerBB yoeB pfyP yoaK
0.5017237 0.5217380 0.5240529 0.6263291 0.6812779 0.6944004 0.7540182
0.7933917
average expression stability M: 0.6371165
gene with lowest stability (largest M value): yoaK
Pairwise variation, ( 7 / 8 ): 0.08989025
#####
Step 4 :
gene expression stability values M:
comEA aepY ymxH mmsA gerBB yoeB pfyP
0.4685845 0.4749793 0.4893698 0.5896670 0.6176832 0.6313060 0.8235834
average expression stability M: 0.5850247
gene with lowest stability (largest M value): pfyP
Pairwise variation, ( 6 / 7 ): 0.1113945
#####
Step 5 :
gene expression stability values M:
aepY comEA ymxH gerBB yoeB mmsA
0.4135145 0.4340826 0.4515760 0.5262518 0.5381601 0.5740228
average expression stability M: 0.4896013
gene with lowest stability (largest M value): mmsA
Pairwise variation, ( 5 / 6 ): 0.0861195
#####
Step 6 :
gene expression stability values M:
aepY comEA ymxH gerBB yoeB
0.3848249 0.4389842 0.4551686 0.4711137 0.4868614
average expression stability M: 0.4473906
gene with lowest stability (largest M value): yoeB
Pairwise variation, ( 4 / 5 ): 0.08268522
#####
Step 7 :
gene expression stability values M:
aepY comEA ymxH gerBB
0.3515844 0.4009611 0.4320285 0.4997325
average expression stability M: 0.4210766
gene with lowest stability (largest M value): gerBB
Pairwise variation, ( 3 / 4 ): 0.115331
#####
Step 8 :
gene expression stability values M:
aepY comEA ymxH
0.3179077 0.3239564 0.3853981
average expression stability M: 0.3424207
gene with lowest stability (largest M value): ymxH
Pairwise variation, ( 2 / 3 ): 0.1211631
#####
Step 9 :
gene expression stability values M:
comEA aepY
0.256466 0.256466
average expression stability M: 0.256466
> ranks
rank DSM319 pSSBm85
1 1 aepY
2 1 comEA
3 2 ymxH
4 3 gerBB
5 4 yoeB
6 5 mmsA
7 6 pfyP
8 7 yoaK
9 8 fdhD
10 9 argR

```

Abbildung 17: Ausgabeprotokoll der Software geNorm zum Ermitteln der stabilsten Referenzgene und das dazugehörige Ranking. Ermittelt wird die durchschnittliche Expressionsstabilität M. Auf diesen Wert basiert die Rangfolge.

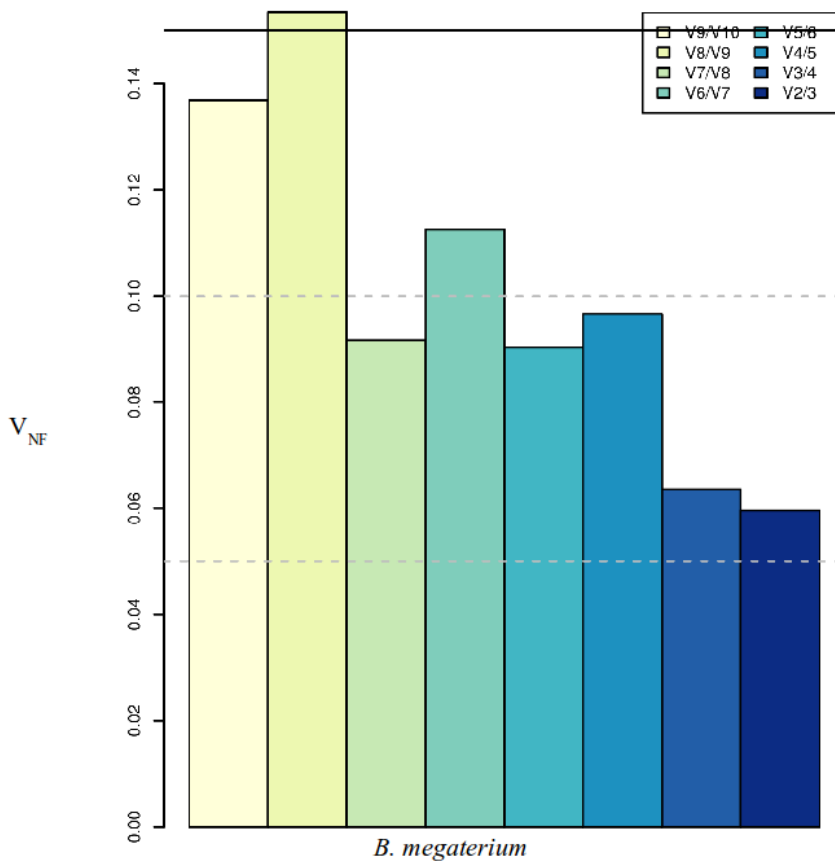


Abbildung 18: Grafische Auswertung der Anzahl der stabilitätsgeordneten Referenzgene mittels geNorm. Bei den zehn getesteten Referenzgenen würden zwei zur Normalisierung ausreichen, da der dritte einen geringen Einfluss besitzt, der vernachlässigt werden kann. Das vorherige Bestimmen des Stabilitätsfaktors M spiegelt die Reihenfolge der Gene wider. Empfohlen wird ein V_{NF} -Wert von $\leq 0,15$. Dieser Schwellenwert ist mit einer waagerechten Linie dargestellt.

3.3.3 Relative Quantifizierung

Die relative Quantifizierung der Transkripte von eGFP mittels qRT-PCR konnte einen ca. 500-fachen Unterschied der hochproduzierenden, grünen Zellen im Vergleich zu niedrigproduzierenden, weißen Zellen aufzeigen, wohingegen die *xylR*-mRNA nur um etwa 150-fach induziert vorlag. Dieser enorme Unterschied lässt den Schluss zu, dass die beobachtete Populationsheterogenität einen immensen Einfluss auf die Produktionsleistung einer Zellkultur hat. Die Expression der *xylT*-RNA blieb unbeeinflusst, obwohl *xylT* Teil des *xyl*-Operons ist (Abb. 19). Jedoch entkoppelt ein Terminator zwischen *xylA* und *xylB* seine Transkription vom P_{xylA} -Promotor und bewirkt dadurch eine niedrigere Transkriptionsrate, was typisch für Membranproteine ist (Schmiedel *et al.*,

1997). Eine positive Rückkopplung durch *xylT* wie im *lac*-Operon beschrieben war demzufolge nicht vorhanden, so dass die Bistabilität in diesem Fall durch einen anderen Mechanismus vermittelt werden muss.

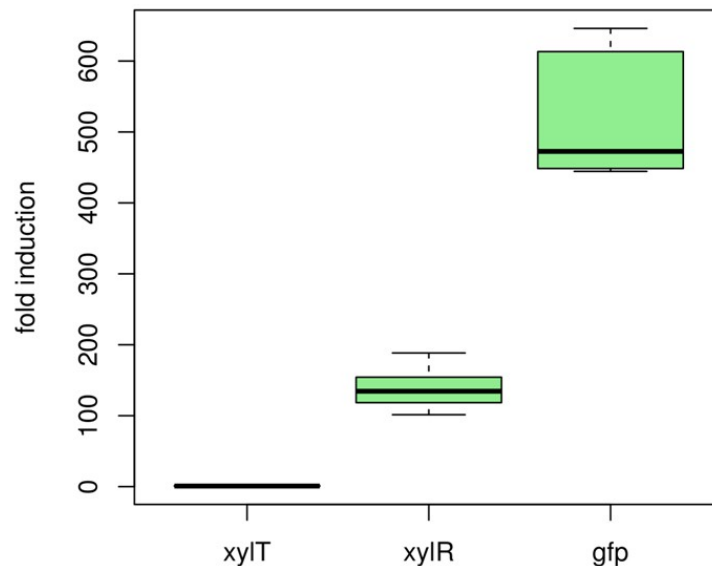


Abbildung 19: Ergebnisse der Echtzeit-PCR in Form eines Boxplots des Produktionsstammes pSSBm85. Die relative Quantifizierung der sortierten Fraktionen zeigte, dass *xylR* und besonders eGFP stark induziert wurden, während die Expression von *xylT* unbeeinflusst blieb.

3.3.4 Plasmid Copy Number (PCN)

Zunächst sollte geklärt werden, ob dieses unterschiedliche Produktionsverhalten an verschiedenen Kopienzahlen der Plasmide beider Subpopulationen festzumachen ist. Würden leistungsstarke Zellen mehr Plasmide besitzen, könnten sie größere Mengen an eGFP produzieren. Um der Frage nachzugehen, wurde die Plasmid Copy Number (PCN) mittels DNA oder über die Anzahl der Zellen von beiden Untergruppen bestimmt.

Die Effizienzberechnung für die PCN erfolgte mittels einer Eichgerade, um inhibitorische Effekte ab einer gewissen Zellzahl ausschließen zu können. Es wurden ausschließlich Eichgeraden verwertet, deren Regressionsgerade einen Korrelationswert größer 0,99 betrug. Der Probenumfang bestand aus sechs verschiedenen DNA-Konzentrationen. Bei dem ausgewählten Beispiel von pSSBm85 2 Std nach Induktion

in Abbildung 20 betrug die Effizienz (E) 1,88 und wurde nach folgender Formel berechnet: $E = 10^{-1/\text{Steigung}}$.

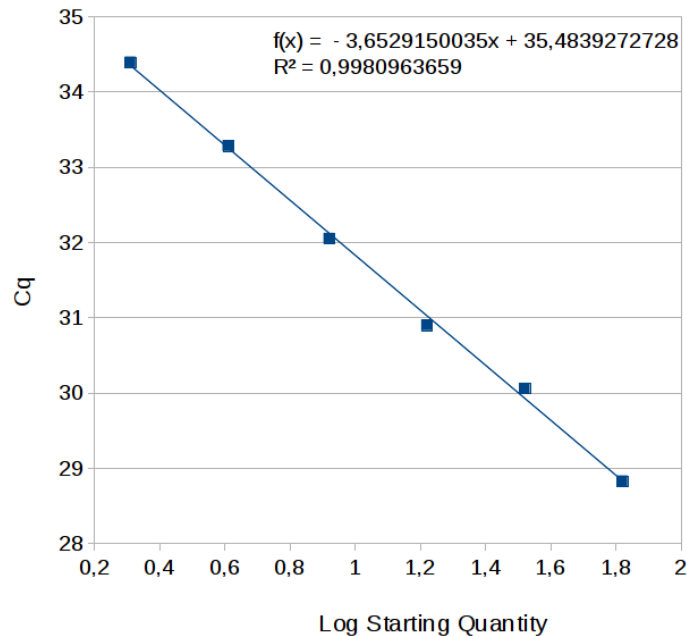


Abbildung 20: Standardgerade einschliesslich der ermittelten Regressionsgerade für eine Effizienzberechnung am Beispiel pSSBm85 2 Std nach Xylosezugabe. Die verwendeten DNA-Konzentrationen wurden logarithmiert und gegen den entsprechenden Cq-Wert aufgetragen. Die Effizienz, die aus der Steigung ermittelt wurde, lag bei 1,88.

Grundsätzlich konnten enorme Schwankungen der PCN bezüglich der biologischen Replikate bestimmt werden, nicht aber innerhalb der technischen Replikate. Die PCN betrug zwischen 260 und 980.

Auf Grund der problematischen Aussagekraft dieser Resultate wurden zur Messung der Plasmidmenge weitere Experimente durchgeführt. Ausgehend vom Vektor pSSBm85 wurde eine translationale mCherry-Fusion mit dem Repressor XylR vorgenommen. Durch Bindung von XylR-mCherry an das Plasmid können sowohl die Verteilung, als auch die Lokalisation der Plasmide visualisiert werden, die Bedeutung der Plasmidkopienzahl untersucht und erste Einblicke in die Segregation von multicopy Plasmiden erlangt werden. Darüber hinaus sollten Lokalisations- und Verteilungsstudien über FISH-Experimente und DAPI Färbungen bestätigt werden (Kapitel 3.5).

3.4 Zeitraffer-Fluoreszenzmikroskopie

Aufgrund seiner enormen Größe und hohen Zellteilungsrate ist *B. megaterium* im Vergleich zu anderen Bakterien besonders gut für die Lebendzellmikroskopie geeignet. Die quantitative Zeitraffer-Fluoreszenzmikroskopie sollte Aufschluss über das Produktionsverhalten der plasmidtragenden Stämme mittels automatisierter, mehrdimensionaler Bildaufnahme geben.

3.4.1 Entstehung von stark- und schwachproduzierenden Zelllinien

Zunächst wurde der Produktionsstamm pSSBm85 untersucht. Bei diesem Film (Münch *et al.*, 2015) begannen vereinzelt in den ersten 2 Stunden nach Induktion ein paar verstreute Zellen eGFP zu produzieren (Abb. 21A). Besonders auffällig war dabei, dass nach der Zellteilung einer eGFP-produzierenden Zelle eine der beiden Tochterzellen weiterhin eGFP produzierte, während die andere Tochterzelle die Produktion einstellte. Darüber hinaus wurde die eGFP-Produktion in der zweiten Tochterzelle nicht nur gestoppt, vielmehr verringerte sich auch nach jeder Zellteilung die Fluoreszenz (Abb. 21B-E). Diese Abnahme kann nicht mit einem Ausbleichen von eGFP erklärt werden, da dieses über mehrere Stunden stabil bleibt (Kap. 3.6.2 bzw. Abb. 41). Darüber hinaus wurden diese Phänomene fast ohne Ausnahme für alle Zellen einer wachsenden Mikrokolonie von *B. megaterium* (Film 1 Anhang 8.1) beobachtet. Als Ergebnis gab es einzelne Zelllinien, die große Mengen an eGFP erzeugten, wohingegen der Rest der Population geringe Mengen produzierte. Diese Beobachtung kann nur durch eine ungleiche Verteilung der zellulären Komponenten, die für die rekombinante Proteinproduktion benötigt werden, erklärt werden. Im späteren Verlauf der Wachstumsphase einer Mikrokolonie erreichte eine zunehmende Anzahl von Zelllinien den Zustand der Hochproduzenten, was schließlich zu einer Verschiebung der Zellkultur hin zur Produktion von eGFP führte.

Mögliche Kandidaten, die nach der Zellteilung solch eine ungleichmäßige Aufteilung von zellulären Komponenten verursachen, könnten neben Komponenten der Proteinbiosynthese entweder das rekombinante Plasmid selbst und / oder der XylR-Repressor sein.

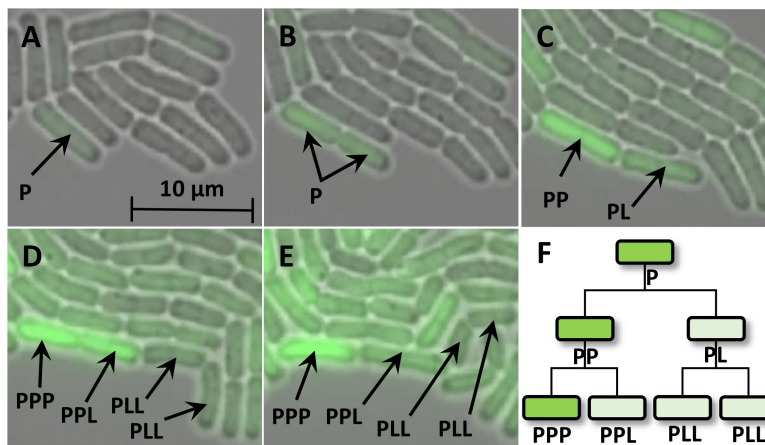


Abbildung 21: Die Bildsequenz eines Zeitraffer-Mikroskopie-Filmes (Münch *et al.*, 2015) zeigen eine wachsende Mikrokolonie von *B. megaterium* mit dem Produktionsplasmids pSSBm85 während der eGFP-Produktion (A - E: 120, 150, 180, 210 und 260 Minuten nach der Induktion der Genexpression durch die Zugabe von Xylose). Produzierende Zellen (P) erzeugten eine Tochterzelle, die eine erhöhte Fluoreszenz aufwies (PP), während die andere kaum fluoreszierte (PL). Der daraus entstehende Zelllinienbaum (F), zeigt schematisch eine Linie von hochproduzierenden Zellen, neben schwachproduzierenden Zellen. Der Film 1 ist Anhang 8.1 zu entnehmen.

3.4.2 Einfluss des Xylose-Repressors auf die eGFP-Produktion

Um zunächst den Einfluss von XylR zu untersuchen, wurde ein Plasmid ohne das *xyIR*-Gen konstruiert (pKMMBm5). Dies erzeugte keine echte Nullmutante, da weiterhin eine Kopie von *xyIR* im Genom DSM 319 vorhanden war. Allerdings wurde dadurch die Anzahl der freien XylR-Proteine in Bezug auf die zur Verfügung stehenden XylR-Bindungsstellen auf dem multicopy Plasmid vernachlässigbar. Auch mit diesem Plasmid konnten Zelllinien von leistungsstarken und leistungsschwachen Produzenten verfolgt werden (Abb. 22A), was praktisch identisch war zum Produktionsplasmid. Eine Batch Kultur von pKMMBm5 im Stamm DSM 319 produzierte in der Abwesenheit von Xylose dauerhaft eGFP (Abb. 23B). Film 4 und 5 (Anhang 8.1) zeigen die Situation vor und nach Induktion. Die beschriebene Kulturheterogenität ließ sich durch eine Δ *xyIR*-Mutante nicht aufheben, was den Schluss nahe legt, dass nicht ausschließlich der Repressor, sondern andere Faktoren maßgeblich an dem Phänomen beteiligt sind. Ebenso bestätigte die Analyse einer genomischen *xyIR*-Mutante (Stamm WH321 von Rygus and Hillen, 1992), die das beschriebene Plasmid pKMMBm5 trug, dass die Heterogenität weiterhin bestehen blieb (Abb. 22B). Da dies eine echte Nullmutante war

ist es naheliegend, dass der XylR-Repressor mit großer Wahrscheinlichkeit nicht die molekulare Ursache der beobachteten Kulturheterogenität ist. Demzufolge sollte der Fokus gezielt auf die Plasmidverteilung gerichtet werden.

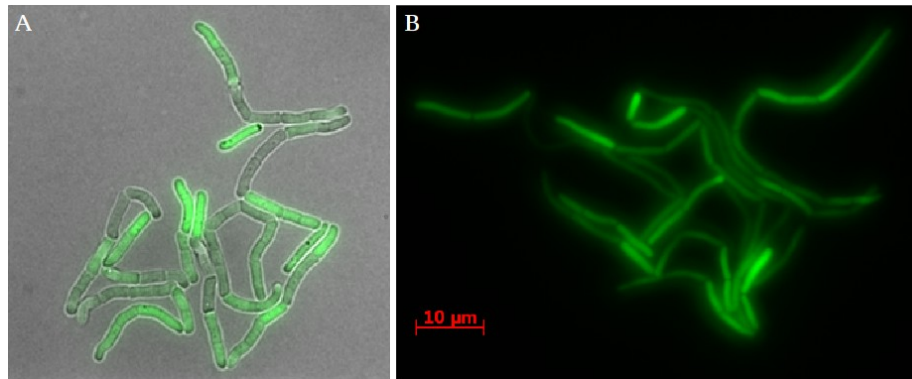


Abbildung 22: Zellheterogenität in Abwesenheit von XylR. Auch bei einem Plasmid ohne das *xylR*-Gen (pKMMBm5) konnte eine extreme Produktionsheterogenität festgestellt werden. A) Die plasmidbezogene XylR-Mutante im Stamm DSM319. B) Das selbe Plasmid im Stamm WH321 (genomische Mutante von *xylR*) zeigte ebenso die beobachtete Heterogenität.

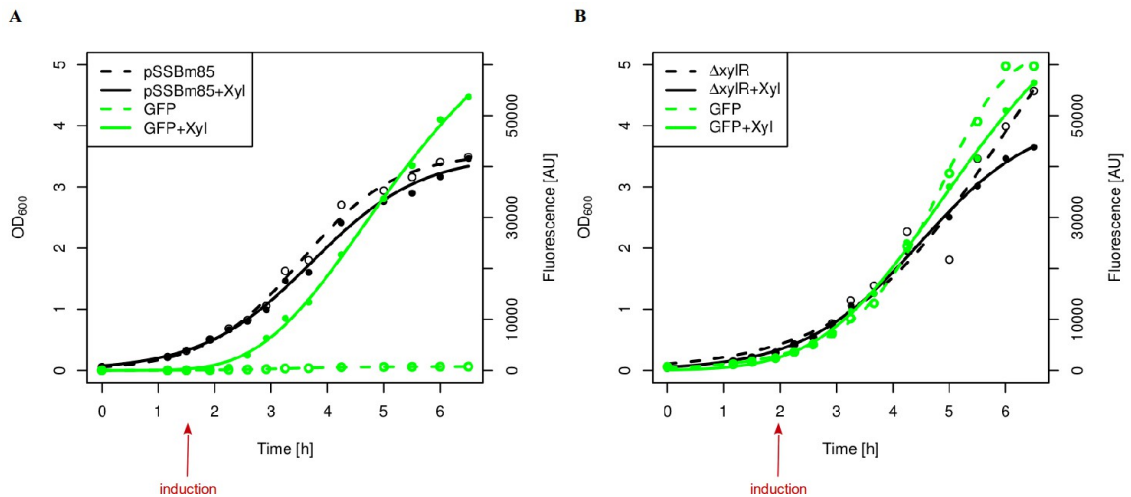


Abbildung 23: Wachstumsverhalten und eGFP-Produktion mittels Batch Kultur des Produktionsstammes pSSBm85 im Vergleich zur plasmidbezogenen XylR-Mutante pKMMBm5. In beiden Fällen konnte kein signifikanter Unterschied bezüglich des Wachstums (schwarze Kurven) ohne Xylosezugabe (gestrichelt) und nach Induktion (durchgezogene Linie) in Erfahrung gebracht werden. A) Erwartungsgemäß wurde kein eGFP vor Induktion des Produktionsstammes pSSBm85 ermittelt (grün gestrichelte Linie), dafür konnte ein rapider Produktionsanstieg (grüne, durchgezogene Linie) nach Xylosezugabe verzeichnet werden. B) Im Gegensatz dazu wurde bei der Mutante, wie zu erwarten, bereits ein eGFP-Anstieg vor Induktion (grün gestrichelte Linie) und nach Induktion (grüne, durchgezogene Linie) registriert.

3.4.3 Zweikanal-Fluoreszenzmikroskopie zur Visualisierung von eGFP und XylR-mCherry

Um die räumliche Verteilung des XylR-Repressors auf Tochterzellen während der Zellteilung zu überwachen, wurde XylR am C-Terminus mit dem rot-fluoreszierenden Protein mCherry fusioniert. Das daraus resultierende rekombinante Plasmid wurde pKMMBm2 bezeichnet. Bei dieser Vorgehensweise war es möglich, die Beziehung zwischen eGFP-Produktion und XylR-Verteilung durch Messen der grünen und roten Fluoreszenz zu bestimmen. Zusätzlich konnte dadurch *in vivo* die Visualisierung und Lokalisierung der Plasmide ermöglicht werden, indem die XylR-mCherry-Fusion an den plasmidkodierten Transkriptionsfaktorbindestellen haftete, ähnlich wie bei einem fluoreszierenden Repressor-Operator-System (FROS).

Zeitraffer-Filme von pKMMBm2 (Film 2 und 3 Anhang 8.1) ergaben ein einzigartiges Verteilungsmuster. Nach der Induktion war die eGFP-Produktion und die Entstehung verschiedener Zelllinien identisch mit der des Produktionsstamms pSSBm85. Darüber hinaus erzeugten einige Zellen rote Spitzen (Abb. 24) was darauf hindeutete, dass sich die XylR-mCherry-Plasmidkomplexe asymmetrisch zu den Zellpolen verteilten, während andere Zellen eine schwache, aber gleichverteilte XylR-mCherry-Fusion aufwiesen. Bei näherer Betrachtung der zeitlichen Abfolge von Aufnahmen wurden diese roten Spitzen häufig an den Randzellen einer wachsenden Zellkette angetroffen (Abb. 24A bis E). Offenbar sammelten sich Plasmidkomplexe mit XylR vorwiegend am alten Zellpol, was die stochastische Plasmidverteilung in neuen Zellpolen durch freie Diffusion in irgendeiner Art und Weise hinderte. Einige dieser rot gefärbten Spitzen wurden während der späteren Wachstumsphase tiefrot (Abb. 24E). Sofern sie kein eGFP besaßen, hörten diese angereicherten Zellen auf sich zu teilen, wohingegen die Nachbarzelle normal weiter wuchs. Diese tiefroten Zellen, die fast ausschließlich am Ende einer Zellkette gefunden wurden sammelten offensichtlich XylR-Komplexe an, während eGFP abgebaut wurde. Bemerkenswert im Hinblick auf die eGFP-Produktionsrate war, dass produzierende Zellen nach der Teilung Tochterzellen generierten, die zwar genetisch identisch waren, bezüglich ihrer Produktion aber zunächst pausierten (Münch *et al.*, 2015). Zum einen könnten sich die eGFP-Moleküle ausdünnen oder zum anderen nach der Zellteilung abgebaut werden. Als Folge darauf

würden einzelne Zelllinien große Mengen an eGFP erzeugen, während die restliche Population kaum oder gar keine Produkte liefern würde. Somit werden höchstwahrscheinlich zelluläre Komponenten, die zur rekombinanten Proteinproduktion benötigt werden, ungleichmäßig auf die Tochterzellen verteilt.

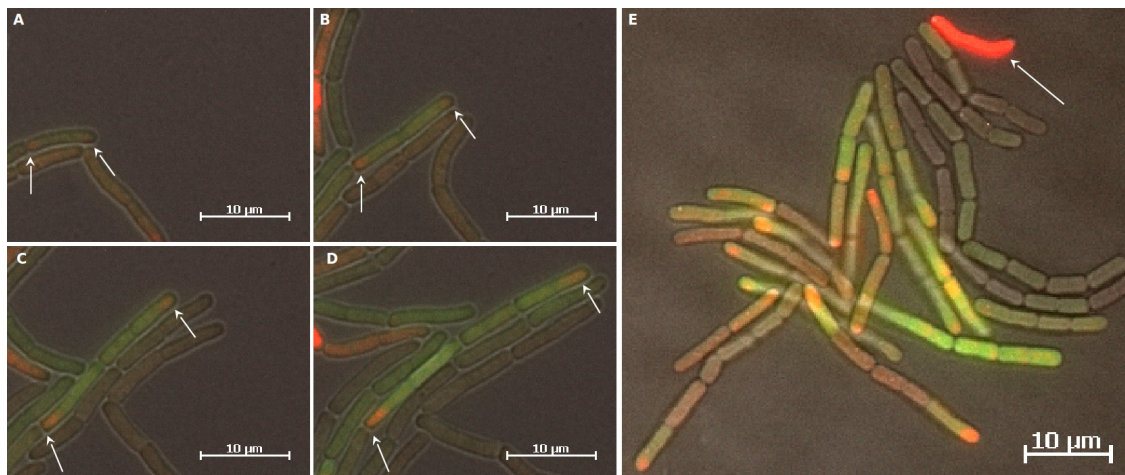


Abbildung 24: Ausschnitte einer wachsenden Kultur von *B. megaterium*, die das Plasmid pKMMBm2 trug zu verschiedenen Zeiten nach der Induktion. In den produzierenden Zellen entwickelten sich rote Spitzen (Pfeile), die an den Zellpolen Komplexe des Plasmids mit XylR-mCherry enthielten (24A-E). Eine schwache Farbintensität korreliert mit einer geringen Produktausbeute und tiefgrüne Zellen gehörten zu den produktionsstarken Zellen. Gelegentlich traten tiefroter Zellen auf, die kein eGFP enthielten und wahrscheinlich XylR-Plasmidkomplexe (Pfeil) ansammelten. E) Diese Zellen stellten ihr Wachstum komplett ein.

Jedoch erreichten im weiteren Verlauf der Wachstumsphase die meisten Zellen den Status der leistungsstarken Produzierer, was schließlich zu einer Verschiebung der Gesamtpopulation gegenüber der Produktion von eGFP führte. Dies war nicht mehr Bestandteil des Films, da dieser bereits aus technischen Gründen nach 4,5 Std endete. Bei längeren Filmen überwuchsen die Zellen das Sichtfeld des Mikroskops bzw. wuchsen dann nicht mehr in einer Schicht als “Monolayer”. Ein Bildausschnitt nach 7 Std zeigte jedoch die in diesem Stadium vorhandene Produktivität (Abb. 25).

Zur Aufklärung der Verteilungsmechanismen von eGFP und XylR-Plasmidkomplexen wurden die Zeitraffer-Filme (Film 2 und 3 Anhang 8.1) des plasmidhaltigen pKMMBm2 Stamms mit Hilfe von Bildanalysemethoden quantifiziert. Dieser Ansatz war aufgrund der Mehrfachkopie des verwendeten Vektors mit nur zwei XylR-Bindestellen pro Plasmid sensitiv genug, visuell gezeigt werden zu können.

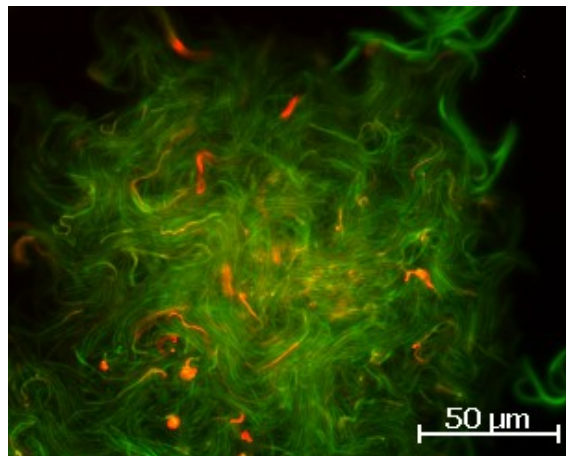


Abbildung 25: Leistungsstarke Produzenten in der späten exponentiellen Phase nach ca. 7 Std. In dieser Phase überwogen die Produzenten, im Vergleich zu den Niedrigproduzierern.

3.4.4 Zelllinienanalyse der Verteilung von eGFP und XylR-mCherry-Plasmid-komplexen

Zum besseren Verstehen der Plasmidweitergabe wurden weitere Analysen einer wachsenden Mikrokolonie des pKMMBm2 Plasmids durch Bildanalyseverfahren von Zeitraffer-Filmen durchgeführt. Dieser Ansatz beinhaltete einen Segmentierungsschritt, in dem Zellen räumlich erfasst werden und einen Trackingschritt, bei dem diese zeitlich verfolgt werden konnten. Zusätzlich wurde die rote und grüne Fluoreszenz (Film 2 Anhang 8.1) bezüglich des Polalters jeder einzelnen Zelle zu jedem Zeitpunkt bestimmt. Das Endergebnis war ein Zelllinienbaum mit detaillierten Informationen über die Dynamik des Zellwachstums, der Plasmidverteilung und der Fluoreszenzintensitäten, was sich durch die eGFP-Produktion und dem Verteilungsverhalten der Plasmidkomplexe mit XylR-mCherry auszeichnete. Darüber hinaus ermöglichte dieser Ansatz sowohl die Bistabilität, als auch das Hystereseverhalten von produzierenden und nichtproduzierenden Zelllinien zu quantifizieren.

Dazu wurde von Zellen mit dem Plasmid pKMMBm2 über 7 Std alle 5 min ein Bildausschnitt aufgenommen. Die Kanäle Hellfeld, eGFP und mCherry wurden zunächst in verschiedenen Graustufen aufgenommen und in definierten Falschfarben dargestellt. In Abbildung 26 ist ein Auszug von Bildausschnitt 47 aus Film 2 (Anhang 8.1) veranschaulicht. Zu sehen ist der Kanal für mCherry (hier rot, Abb. 26A) und eGFP (hier

grün, Abb. 26B) sowie der Hellfeldkanal (Abb. 26C). Die entsprechenden Mischfarben wurden durch Aufeinanderlegen der Kanäle erzeugt (Abb. 26D). Auffallend dabei waren abermals die intensiven Rotfärbungen an den Spitzen von Zellenden (Abb. 27). Der Film begann mit einem Ausschnitt von drei Zellen (Abb. 28A) und endete nach 4,5 Std mit dem Hellfeldbild aus Abbildung 28F. Die Segmentierung einzelner Bildausschnitte via TLM-Tracker zeigten die Bildsequenzen in Abbildung 28.

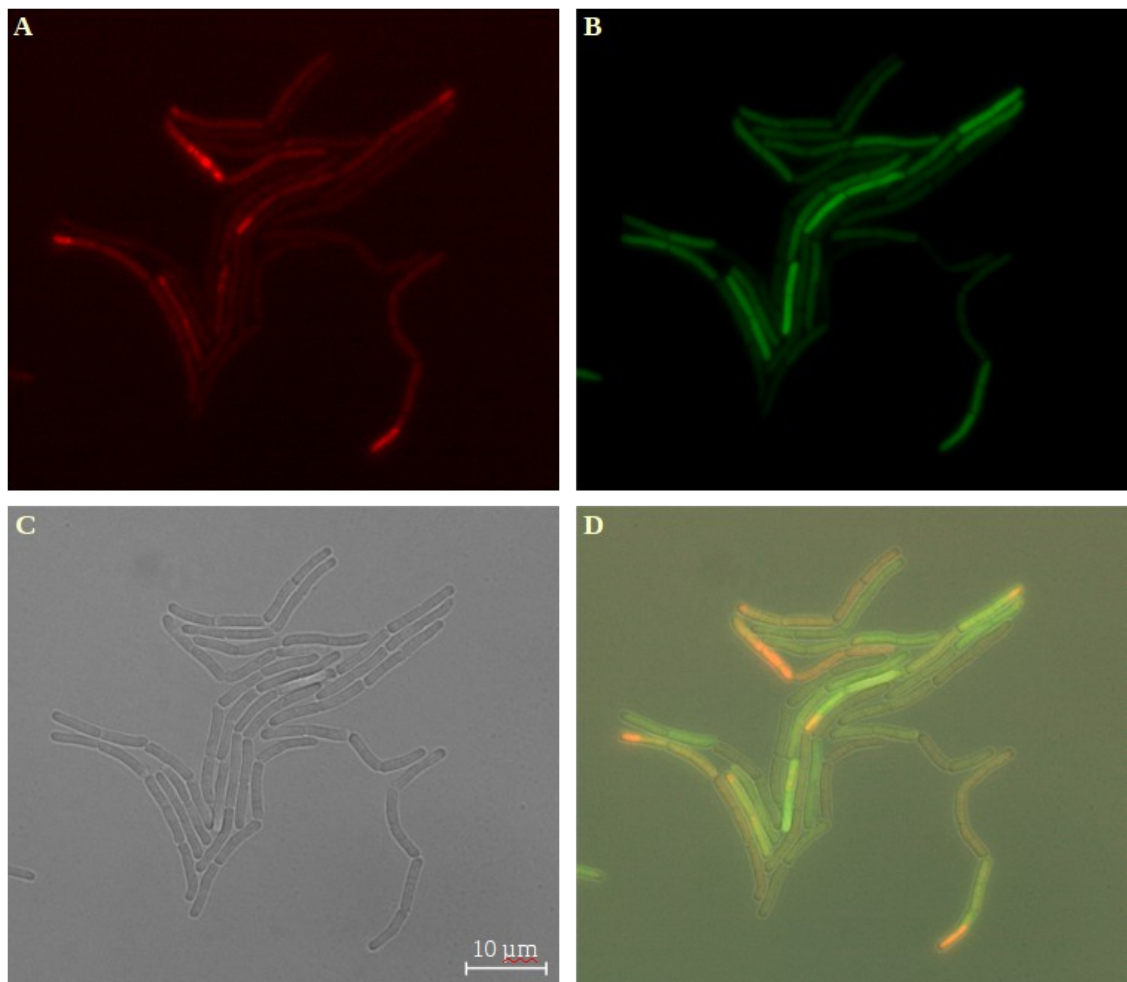


Abbildung 26: Auszug einer mehrdimensionalen Bildaufnahme (Film 2 Anhang 8.1) am Beispiel von Bildausschnitt 47. Dargestellt sind die verschiedenen Kanäle. A) Der Kanal für mCherry wird durch die Falschfarbe Rot dargestellt. B) Der für eGFP durch Grün. C) Im Hellfeldkanal lassen sich die Zellbegren-zungen deutlich erkennen. D) Legt man die Bilder dieser Kanäle aufeinander, erhält man ein Mischbild.

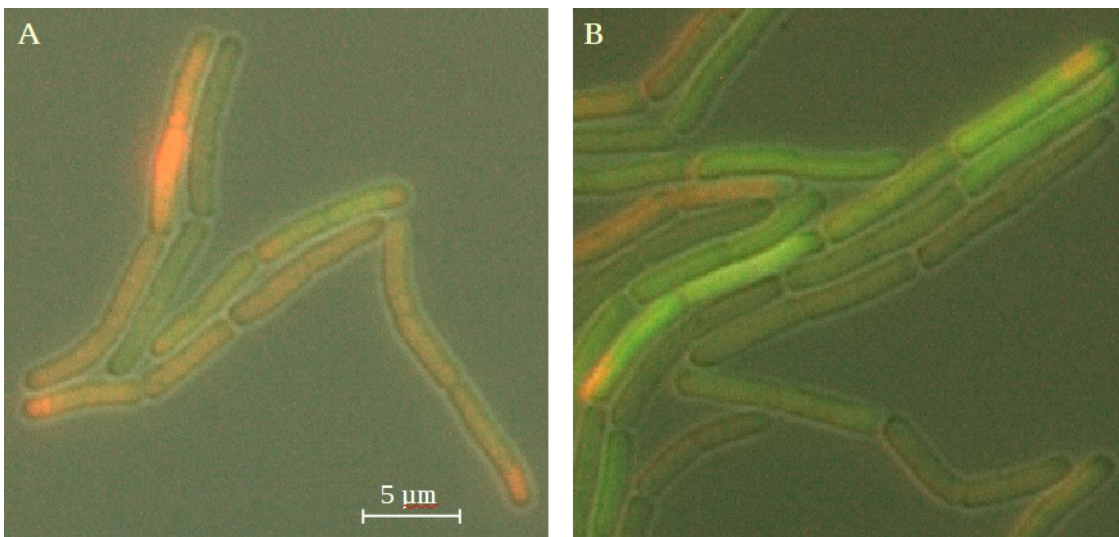


Abbildung 27: Entstehung rot gefärbter Spitzen nach Induktion mit Xylose. A) Bereits zu Beginn einer wachsenden Kultur (1 Std nach Induktion) tauchten rote Spitzen an den Zellpolen auf. B) Sie blieben auch während des Wachstums bestehen (3 Std nach Induktion).

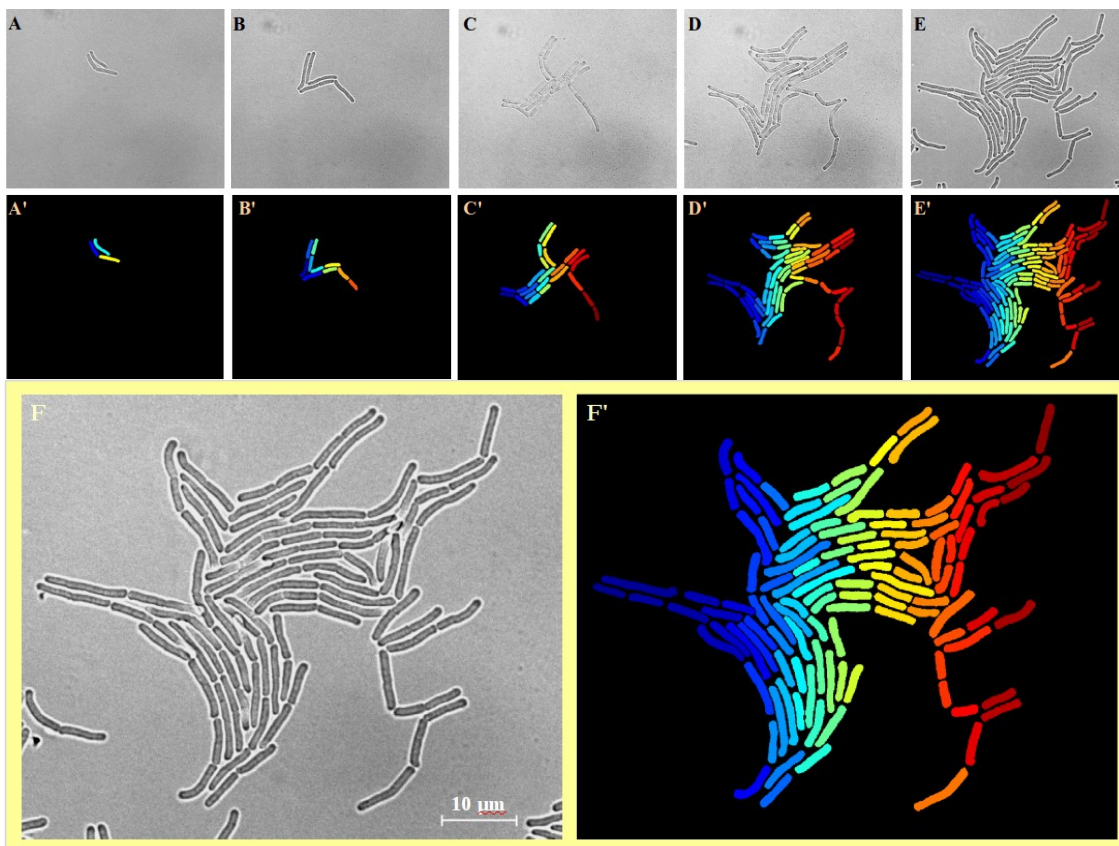


Abbildung 28: Zellsegmentierung des Zeitraffer-Filmes. A bis F zeigt die wachsende Population im Hellfeldkanal und die dazugehörigen mit TLM-Tracker segmentierten Abbilder A' bis F'. F) Mit diesem Bildausschnitt endete der Zeitraffer-Film (Film 2 Anhang 8.1).

Sowohl einige Ausschnitte des Films (Film 2 Anhang 8.1) die im Hellfeldkanal aufgenommen wurden, als auch die dazugehörigen Segmentierungen sind im Folgenden veranschaulicht. Dabei zeigen die Abbildungen A bis F jeweils die sich teilenden Zellen am Fluoreszenzmikroskop im Hellfeldkanal und die entsprechenden Abbilder A' bis F' die Segmentierung nach der Bildanalyse.

Das Tracking ermöglichte die zeitliche Zuordnung von Zelllinienbäumen für eGFP und XylR-mCherry (Abb. 29). Dabei gilt, je höher die Farbintensität, desto höher die Fluoreszenz und umso mehr des entsprechenden Produkts wurde gebildet. Der Zelllinienbaum bestätigte die mikroskopischen Beobachtungen (Abb. 24).

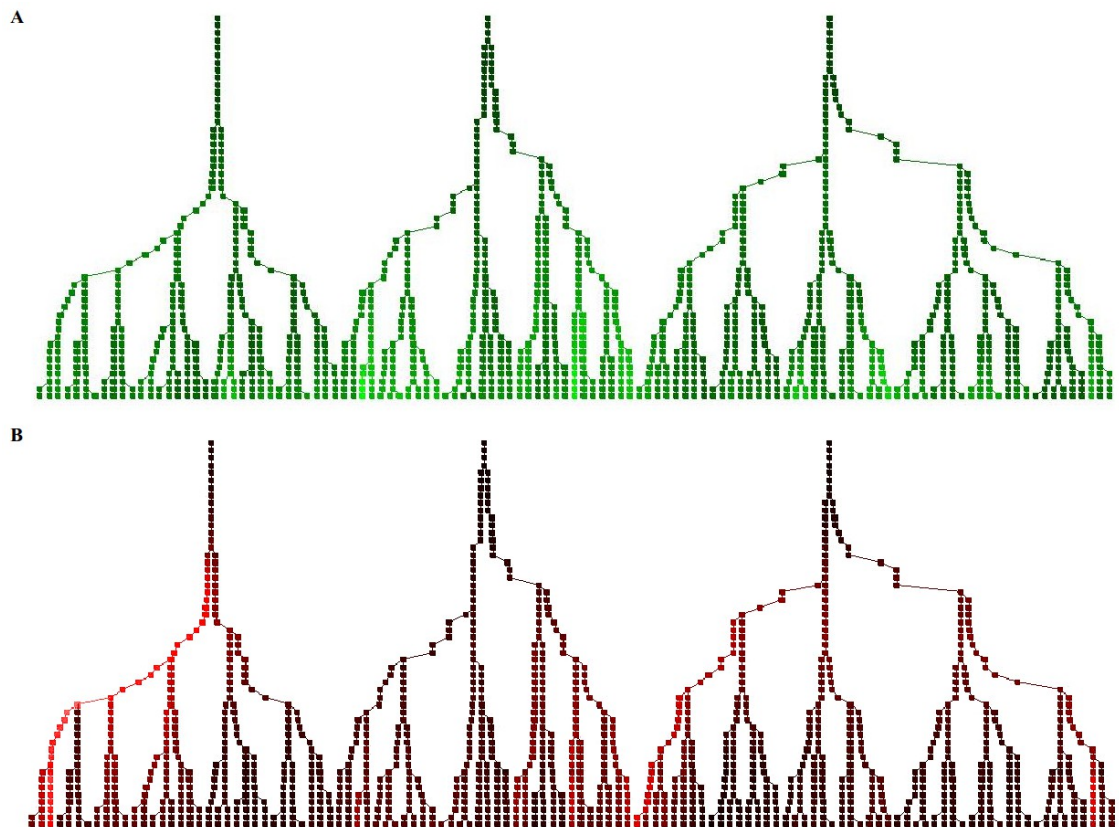


Abbildung 29: Zelllinienstammbäume von pKMMBm2. Ausgehend von drei Bakterienzellen wurde alle 5 min ein Bildausschnitt aufgenommen. Jeder Punkt steht dabei für eine neue Aufnahme, eine Verzweigung ist mit einer Zellteilung gleichzusetzen. A) Der grüne Stammbaum beschreibt das Produktionsverhalten von eGFP. B) Der deckungsgleiche rote Stammbaum zeigt analog die Konzentrationen von XylR-mCherry. Die Intensität der Farbe korreliert mit der Produktmenge. Je intensiver die Farbe, desto höher das Produktionsverhalten dieser Zelle.

Der grüne Baum, der die eGFP-Produktion darstellt, zeigte bei den meisten Verzweigungen eine deutliche Aufspaltung in eine hoch- und eine niedrigproduzierende Zelllinie. Im deckungsgleich roten Baum, der die XylR-mCherry-Konzentration beschreibt, waren nur in vereinzelten Zelllinien starke Signale zu erkennen. Generell war das XylR-mCherry-Signal im Vergleich zu eGFP deutlich schwächer, was durch den stärkeren P_{xylA} -Promotor erklärbar ist und über die qPCR-Daten bestätigt wird. Zudem war auffällig, dass zwischen den Signalen von eGFP und XylR-mCherry keinerlei signifikante positive oder negative Korrelation erkennbar war. Der Datensatz des Zelllinienbaums wurde im nächsten Schritt hinsichtlich der asymmetrischen Verteilung der Fluoreszenzsignale untersucht.

Dazu wurden Boxplots über die Verteilung der Fluoreszenzintensitäten pro Zelle erstellt (Abb. 30), um die Einzelzelldynamik zu erfassen. Es konnte ein doppelt sigmoider Produktionsverlauf von eGFP ermittelt werden mit einer Plateau-Phase zwischen 2 bis 3,5 Stunden. Auffällig war, dass die Streuung der Fluoreszenzintensitäten pro Zelle im Zeitverlauf deutlich zunahm. Der Beginn des XylR-mCherry-Signals gestaltete sich ähnlich. Es konnte ein stetiges Ansteigen bis zu einem Maximum nach 2,5 Stunden gemessen werden. Nach dieser Phase war dann ein kontinuierliches Abfallen zu sehen, was den Schluss zulässt, dass viel mCherry sich in wenigen Zellen sammelt und gleichzeitig weniger neues produziert wird bzw. rote Zellen sich weniger teilen (siehe auch tiefrote Zelle in Abb. 24E).

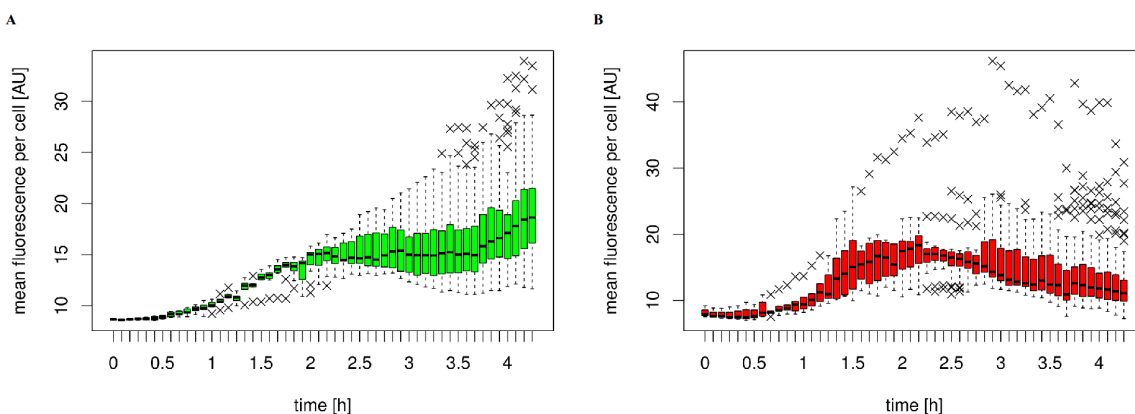


Abbildung 30: Verteilung der Fluoreszenzintensitäten pro Zelle. A) Grüne Boxplots entsprechen der Verteilung von eGFP pro Zelle. B) Rote Boxplots von mCherry.

Ausgehend von der Annahme, dass in heterogenen Kulturen eine ungleichmäßige Ver-

teilung von Plasmiden vorliegt, wurden im Folgenden das asymmetrische Verteilungsverhalten der Plasmidkomplexe über XylR-mCherry untersucht (Abb. 31).

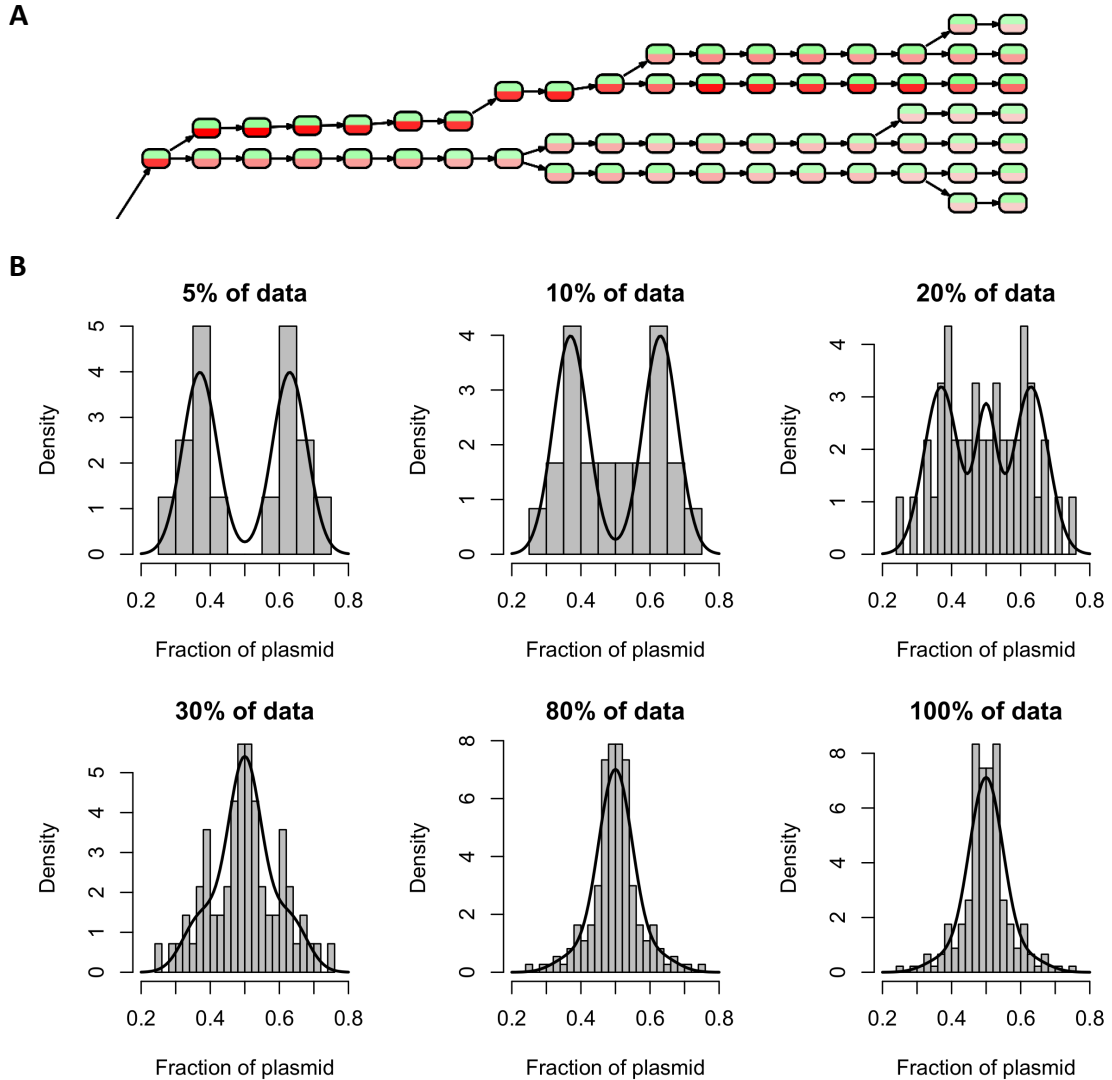


Abbildung 31: Analyse der asymmetrischen Verteilung der XylR-mCherry-Komplexe. A) Ausschnitt des rekonstruierten Zelllinienbaums aus dem Zeitraffer-Film von *B. megaterium* mit dem Plasmid pKMMBm2 (Film 2 Anhang 8.1). Die gefüllten Knoten stellen einzelne Zellen zu bestimmten Zeitpunkten dar. Rote und grüne Färbungen stehen für die jeweiligen Fluoreszenzintensitäten. Hier ist eGFP grün dargestellt und XylR-mCherry rot. B) Zusammenhang der bimodalen Plasmidverteilung zur Plasmidmenge pro Zelle. Die Histogramme zeigen den Anteil der segregierten Plasmide in Bezug auf das Verhältnis jener Zellen, mit dem stärksten XylR-mCherry Signal. Die Zellen wurden nach roter Fluoreszenzintensität sortiert und jeweils auf Daten der stärksten Signale selektiert (5, 10, 20, 30, 80 %). Eine bimodale Verteilung wurde bei 5 % - 20 % der Zellen gefunden, die alle große Mengen an XylR-mCherry beinhalten. Die gefitteten Kurven wurden durch ein heterogenes Modell generiert (Münch *et al.*, 2015). Stark fluoreszierende Zellen wurden durch eine binormale Verteilung durch eine Überlagerung von zwei Normalverteilungen, die bei $0,5 \pm 0,13$ lokalisiert war (die Varianz lag bei 0,05) gefittet. Die mit geringer Fluoreszenz verbleibenden Mutterzellen wurden mit einer Normalverteilung (mit Mittelwert 0,5 und Varianz 0,03) gefittet.

Ein Ausschnitt des Baums stellt in grün die eGFP- und in rot die mCherry-Produktion vereint in einer Grafik dar – der komplette Zelllinienbaum ist Abb. 32 zu entnehmen – und verdeutlicht nochmals den über Zeitraffer-Filmen gemessenen Effekt. Dabei wurde das Verhältnis der vererbten Plasmide jeder einzelnen Tochterzelle durch die Summe der Fluoreszenzintensität beider Tochterzellen dividiert. Daraus ergibt sich ein Wert von 0,5 bei einer Gleichverteilung, wohingegen deutlich höhere oder niedrigere Werten eine asymmetrische Verteilung reflektieren. Während sich für die Gesamtpopulation eine Normalverteilung ergab, zeigten stark fluoreszierende Zellen eine bimodale Verteilung, die mit dem Hartigan's Dip Test als signifikant bestätigt wurde. Diese Ergebnisse bestätigten die Beobachtung, dass asymmetrische rote Spitzen nur in Gegenwart von starken XylR-mCherry-Signalen erzeugt werden. War die Konzentration dieses XylR-mCherry-Komplexes niedrig, erhielten die beiden Tochterzellen normalerweise annähernd die gleiche Menge. Bei höheren Konzentrationen führte eine ungleiche Verteilung zu erheblichen Unterschieden in den Tochterzellen. Die ungleich verteilten Komplexe wurden gleichermaßen in den alten und neuen Zellpolen gefunden. Somit schien das beschriebene Phänomen stochastischer Natur zu sein, und stand nicht im Zusammenhang zum älteren Zellpol. Das asymmetrische Verteilungsverhalten der XylR-Plasmidkomplexe lieferte demzufolge keinen Hinweis auf die bestehende Zellalterung. Ferner war der Versuch zur Kolo-kalisation vermutlich nicht empfindlich genug, um einzelne verstreut liegende Plasmide innerhalb der Zelle sichtbar zu machen, was im nächsten Schritt mittels FISH-Experimenten analysiert werden sollte.

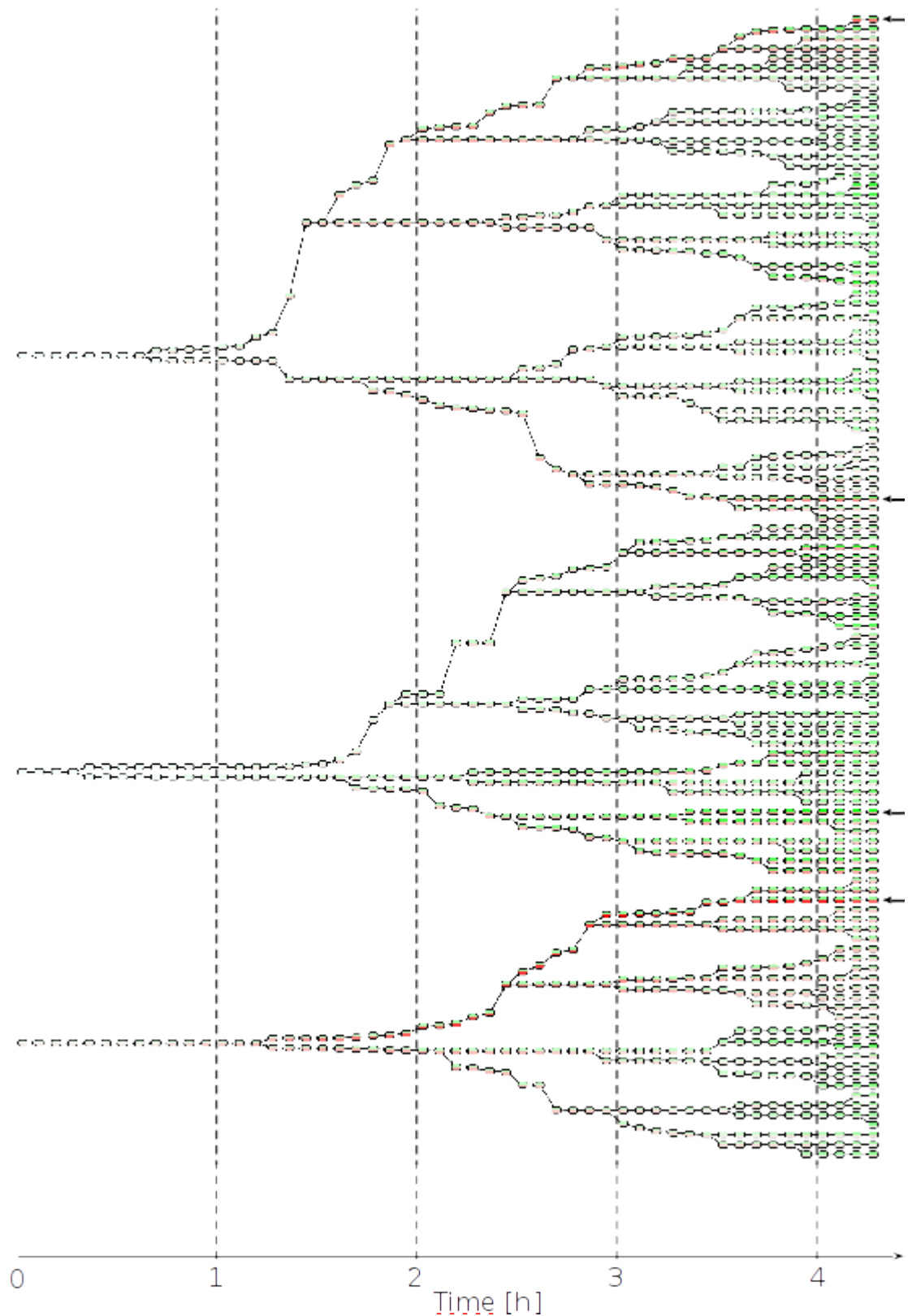


Abb. 32: Rekonstruierter Zelllinienbaum aus dem Zeitraffer-Film von *B. megaterium* mit dem Plasmid pKMMBm2 (Film 2 Anhang 8.1). Die Signale für XylR-mCherry und eGFP wurden in einem Baum zusammenfassend dargestellt. Insgesamt wurden aus 4,5 Stunden Filmmaterial 52 Einzelbilder verarbeitet, die 231 verschiedene Zelllinien und 1326 analysierte Zellzustände reflektierten. Die darge-

stellten Knoten stehen für einzelne Zellen innerhalb der verschiedenen Zeitpunkte. Rote und grüne Färbungen der Knoten stellen die Fluoreszenzintensitäten für die einzelnen Zellen dar. Die Pfeile zeigen die Zelllinien mit starkem rotem Fluoreszenzsignal was darauf hindeutet, dass in diesen Zellen große Mengen an Plasmid-XylR-mCherry-Komplexe angesammelt wurden.

3.5 Plasmidlokalisierung mittels Fluoreszenz *in situ* Hybridisierung (FISH)

Weil mit Lokalisierungsproblemen eines mCherry-Fusionsproteins zu rechnen ist (Landgraf *et al.*, 2012), wurden die Ergebnisse der asymmetrischen Plasmidverteilung über einen unabhängigen Ansatz mittels FISH verifiziert. Um den Einfluss von Proteinproduktion und Plasmidverteilung zu visualisieren, wurden für den direkten Nachweis zellulärer Plasmide FISH-Experimente mit dem Produktionsstamm pSSBm85 durchgeführt (Abb. 33A). Die verdaute, mit Cy3-dCTP markierte Plasmid-DNA diente dabei als Sonde. Die Ergebnisse bestätigen eindeutig die ungleichmäßige Verteilung nach der Induktion. Zusätzlich zeigten einige Zellen ein extrem starkes Fluoreszenzsignal und bekräftigten dadurch die Akkumulation von Plasmiden in diesen Zellen, was bereits in Zeitraffer-Filmen nachgewiesen wurde (Film 2 Anhang 8.1 und Abb. 24). Im Gegensatz dazu wurden rekombinante Zellen von *B. megaterium* ohne Zusatz von Xylose kultiviert. Die ungleiche Verteilung von Plasmiden fehlte hier weitgehend (Abb. 33B), was die Schlussfolgerung nahelegt, dass die Proteinbiosynthese starken Einfluss auf die Verteilungsmechanismus der Plasmide ausübt. Als weitere Kontrolle wurden FISH-Analysen von *B. megaterium* vorgenommen, die das Plasmid pBC16 trugen (Abbildung 33C). Dieser Vektor (Abb. 4) trägt kein eGFP, besitzt eine Größe von 4,6 kB (Bernhard *et al.*, 1978) und ist der Vorläufer des 7,2 kB großen Produktionsplasmids pSSBm85 (Stammen *et al.*, 2010). Auch in diesem Fall konnte ausschließlich eine gleichmäßige Verteilung beobachtet werden, die als getrennte Focigruppen über die ganze Zelle verteilt waren. Zusätzlich wurde dieses Plasmid in *Bacillus subtilis* getestet und erzielte ein ähnliches Resultat (Abb. 33D). In vorangegangenen Studien wurde gezeigt, dass highcopy Plasmide in der Regel in Form von Komplexen in den Zellen vorliegen, was in FISH- oder FROS-Experimenten als fluoreszierende Foci erkennbar ist. Somit bestätigten die vorliegenden Ergebnisse diese Arbeiten und zeigen, dass kein Artefakt vorlag.

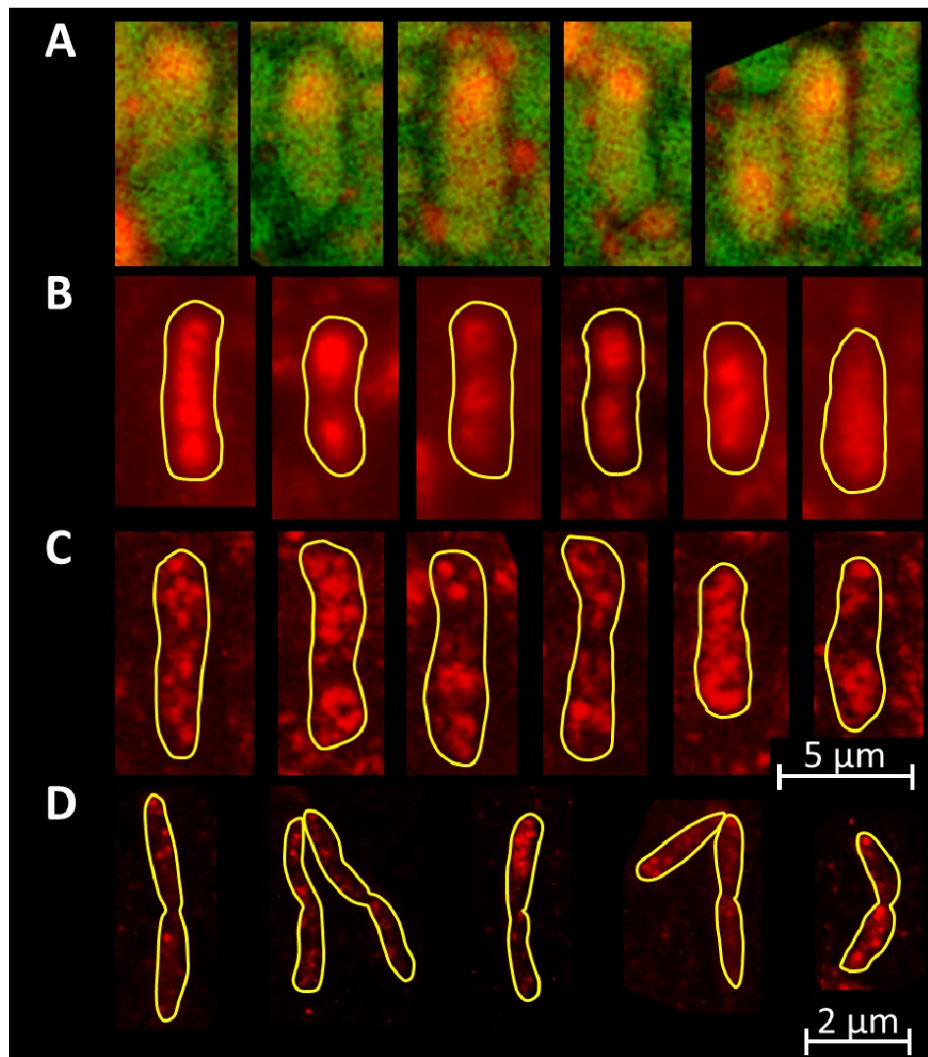


Abbildung 33: FISH-Analysen zur direkten Visualisierung der Plasmide. A) In mit Xylose induzierten Zellen zeigte das Plasmid pSSBm85 eine ungleiche Verteilung und blieb bevorzugt am Zellpol. B) Ohne Geninduktion wurde hingegen die Verteilung dieses Plasmids kaum beeinflusst. C) Das kleinere Vorläuferplasmid pBC16 wurde in mehreren kleineren Foci über die ganze Zelle verteilt lokalisiert. D) Das gleiche Muster und dieselben distinkten Foci wurde mit diesem Plasmid in *B. subtilis* gefunden. Die weißen Linien kennzeichnen die Zellgrenzen.

Zunächst wurde analysiert, ob die Plasmidlokalisierung von der chromosomalen DNA des Nucleoids beeinflusst wird. Bei kleineren Bakterien wie *E. coli* wurde gezeigt, dass highcopy Plasmide den nucleoidfreien Raum innerhalb der Bakterienzelle einnehmen (Gordon *et al.*, 1997). Dieser Sachverhalt wurde mit pKMMBm2 überprüft, da dieses Plasmid bereits nach Induktion mittels roter Fluoreszenz visualisiert werden kann (Abb. 34A) und gleichzeitig die genomische DNA mit der blauen Fluoreszenz von DAPI

anfärbbar ist (Abb. 34B). Von beiden Fluoreszenzkanälen wurde ein Mischbild generiert (Abb. 34C). Daraufhin wurden verschiedene Stadien der Zellteilung untersucht. Vor einer Zellteilung waren die Nucleoide im zentralen Teil der Zelle lokalisiert, während die XylR-mCherry-Plasmidkomplexe vor allem über die ganze Zelle verteilt waren, einschließlich ihrer Zellpole (Abb. 34, Reihe 1). Im Zuge der Zellteilung, wurden Plasmidkomplexe dann an einem der Tochterzellpole lokalisiert, während die Nucleoide in unmittelbarer Nähe der neuen Pole anzutreffen waren (Abb. 34, Reihe 2). Schließlich war die gesamte Zelle nach ihrer Teilung mit Plasmid-DNA besetzt, während sich die Nucleoid-DNA auf den mittleren Teil beschränkte. Durch eine asymmetrische Plasmidverteilung ergaben sich terminale Zellen einer Bakterienkette, die hohe Mengen an XylR-Plasmidkomplexen enthielten (Abb. 34C, Reihe 3).

Erwähnenswert ist, dass eher große Plasmidkomplexe an den Polen detektiert werden konnten, im Vergleich zu den verteilten distinkten Foci des Plasmids pBC16. Verschiedene Studien in *B. subtilis* haben gezeigt, dass sich aktive Ribosomen an den Polen der Zelle befinden (Sojka *et al.*, 2007) und die zelluläre Lokalisation von DNA und Ribosomen überlappen, was auf einen gekoppelten Transkriptions- und Translationsprozess (Mascarenhas, 2001) hinweist.

Dennoch ist trotz der Bildung von Plasmidclustern an den Polen gewährleistet, dass mindestens ein Plasmid nach der Zellteilung der neuen Zelle zugeordnet wird, obwohl diese vorzugsweise in der Nähe der Pole als Gruppen lokalisiert sind (Pogliano, 2001 und Gordon, 1997). Im allgemeinen wird angenommen, dass highcopy Plasmiden durch eine zufällige Diffusion aufgeteilt werden, während lowcopy Plasmide über ein Partitionierungssystem verteilt werden, das in Zentrumsnähe liegt und sich ähnlich wie Minichromosomen verhalten (Meinhardt *et al.*, 1989). Im vorliegenden Fall jedoch scheint der Hauptgrund für die ungleichmäßige Plasmidverteilung die polare Fixierung während der rekombinanten Proteinproduktion zu sein, was im Widerspruch zur gegenwärtig angenommenen Meinung der freien Diffusion von multicopy Plasmiden steht.

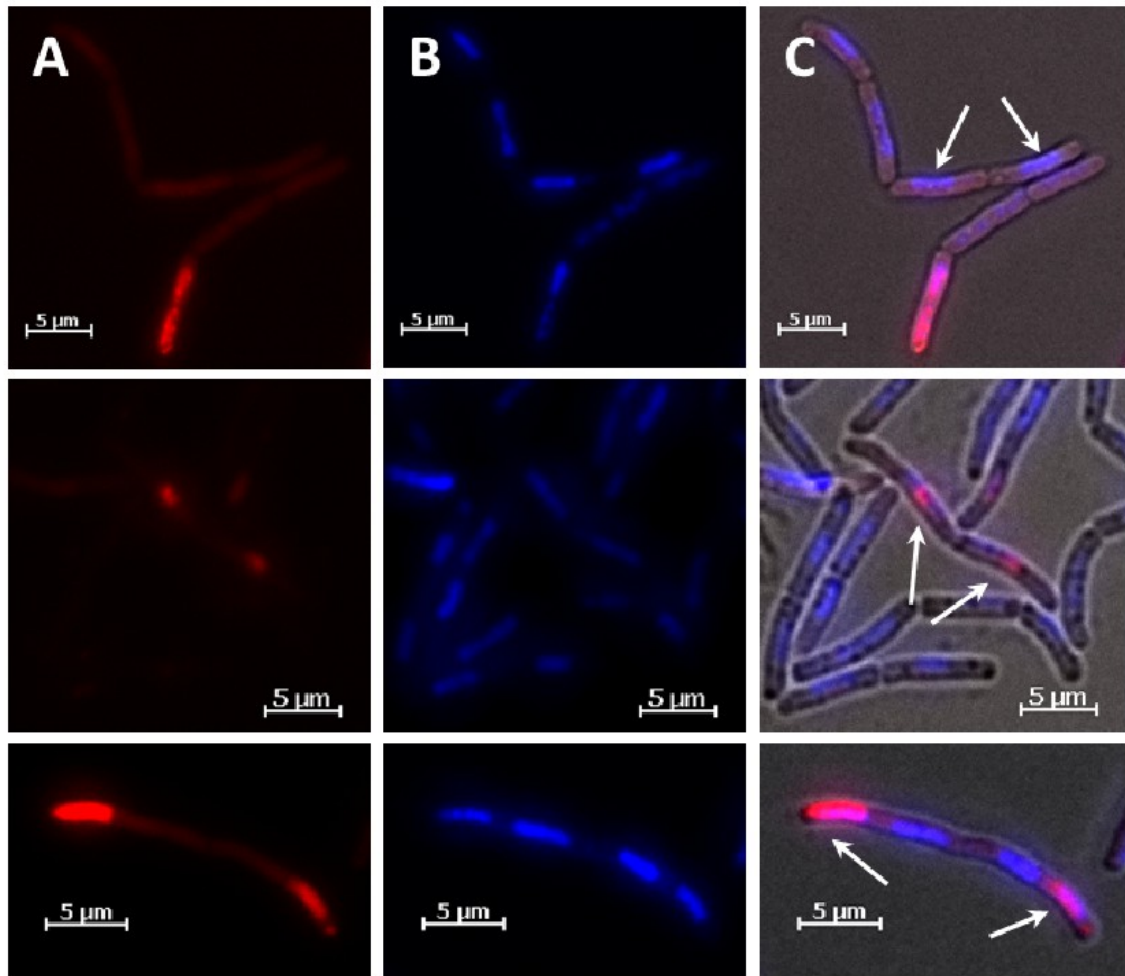


Abbildung 34: Lokalisation des Plasmids pKMMBm2 und seiner chromosomalen DNA zu verschiedenen Phasen der Zellteilung. Die Plasmide sind auf Grund der Komplexbildung mit der XylR-mCherry-Fusion rot dargestellt (A), wohingegen in blau (B) die mit DAPI angefärbte genomische DNA zu sehen ist. Die beiden Fluoreszenzbilder wurden zusammen mit dem Bild des Hellfeldkanals vereint und durch ein Mischbild (C) dargestellt. Vor der Zellteilung wurden Nucleoide im zentralen Teil der Zelle lokalisiert, während die Pole frei blieben (Pfeile in Zeile 1 bei Spalte C). Das Fusionsprotein XylR-mCherry war überwiegend in der ganzen Zelle verteilt. Während der Zellteilung wurden die Plasmid-XylR-mCherry-Komplexe an den alten Polen lokalisiert (Pfeile in Zeile 2, Spalte C), wohinegen die Nucleoide in der Nähe der neuen Pole vorlagen. Dies führte zu Bakterienketten mit terminalen Zellen, die hohe Mengen des Plasmidkomplexes mit XylR-mCherry (Pfeile in Zeile 3, Spalte C) beinhalten.

3.6 Durchflusszytometrie

Die vorliegenden Daten deuten darauf hin, dass die Entstehung der beobachteten Populationsheterogenität ein kontinuierlicher Prozess ist. Um die Dynamik dieses Phänomens zu verstehen, wurden wachsende Kulturen von *B. megaterium* mittels

Durchflusszytometrie auf Einzelzellebene untersucht. Zunächst wurde durch die Zugabe von Xylose in der frühen exponentiellen Wachstumsphase bei Stämmen mit unterschiedlichen Plasmiden eine rekombinante Genexpression induziert und sowohl die eGFP-Produktion, als auch wenn möglich die Produktion von XylR-mCherry zu verschiedenen Zeitpunkten mit Hilfe der Durchflusszytometrie analysiert. Als Vergleich dienten die jeweils uninduzierten Stämme.

Grundsätzlich wurde bei einer OD_{578} von 0,2 induziert sowie 1, 2, 3 4, 5 und mitunter 6 Std nach Zugabe von Xylose eine Analyse mit dem Durchflusszytometer durchgeführt. Dargestellt wurden FSC bzw. der rote Fluoreszenzkanal (FL2) gegen eGFP. In allen Experimenten wurden jeweils 100.000 Ereignisse gemessen.

3.6.1 Dynamik der Populationsheterogenität während der heterologen Proteinproduktion von eGFP

Ein großer Teil der Zellen (ca. 72 %), die das Plasmid pSSBm85 enthielten zeigte bereits vor Induktion einen sehr niedrigen eGFP-Level, was auf eine basale Expressionsrate durch den P_{xylA} -Promotor in Abwesenheit von Xylose hindeutet (Abb. 35A). Bereits der Wildtyp besitzt eine minimale Eigenfluoreszenz, die es zu berücksichtigen galt (siehe auch Abb. 36A). Mit der Zeit begann ein wachsender Anteil an Zellen, große Mengen an eGFP zu produzieren, wobei sich dabei jedoch ein signifikanter Prozentsatz einer niedrigproduzierenden Subpopulation herauskristallisierte. Anfänglich konnten 29,1 % (eine Stunde nach der Induktion), danach 16,9 % (zwei Stunden nach der Induktion) und 11,1 % (drei Stunden nach Induktion) sowie anschließend 3,6 % (vier Stunden nach Induktion) einer Population nachgewiesen werden, die jeweils keine oder nur geringe Mengen an eGFP herstellte (Abb. 35B bis E). Fünf Stunden nach der Zugabe von Xylose und einhergehend mit dem Eintritt in die stationäre Wachstumsphase, wechselte ein großer Teil der schwachproduzierenden Zellen in einen produktiven Zustand (Abb. 35F), bis schließlich am Ende der stationären Phase (6 Std nach Induktion) die rekombinante eGFP-Produktionsrate wieder reduziert und zurückgebildet wurde in Richtung des beginnenden Produktionsmusters vor der Induktion mit Xylose (Abb. 35A

und G). Als zusätzliche Probe wurde eine Kultur in der stationären Phase ohne Antibiotikum aufgetragen (35H). Eine deutliche Heterogenität während der Proteinproduktion im plasmidbasierten, rekombinanten Produktionssystem von eGFP in *B. megaterium* konnte somit belegt werden. Ähnliche Resultate wurden unter Verwendung von verwandten Plasmiden erzeugt, die ebenso mit Hilfe der Durchflusszytometrie auf ihre Heterogenität überprüft wurden. Jedoch konnten durch gezielte Veränderungen von dem Expressionssystem beteiligten Komponenten auch signifikant abweichende Populationsverteilungen generiert werden. Erwartungsgemäß besaß der Wildtypstamm DSM 319 (Abb. 36) keine ausgeprägte Heterogenität, wobei gegen Ende der Wachstumsphase das Signal für eGFP leicht erhöht war. Dies deckte sich mit den Zeitraffer-Filmen. Bei einer alternden Kultur zeigten die Septen zum Ende eine verstärkte Eigenfluoreszenz im grünen Bereich (Abb. 36F).

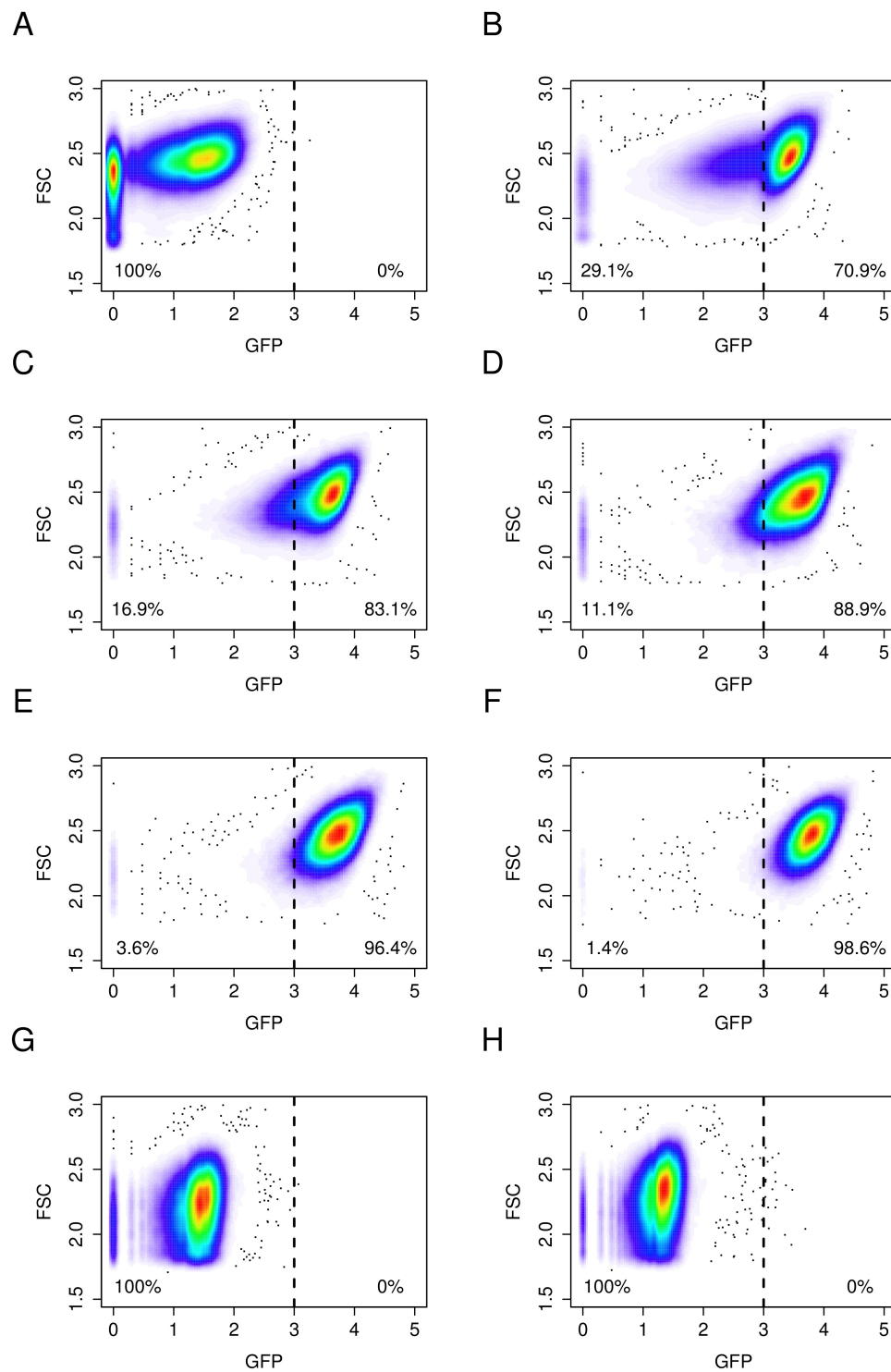


Abbildung 35: Durchflusszytometrische Daten vom Produktionsstamm pSSBm85. Die Dichteplots zeigen FSC aufgetragen gegen das Signal von eGFP im zeitlichen Verlauf nach der Induktion mit Xylose. A) Vor der Induktion und B) – G) 1 bis 6 Stunden nach der Induktion, sowie H) eine Kultur in der stationären Phase ohne Tetracyclin. Die gestrichelte Linie markiert die Auftrennung bezüglich der Produktivität in zwei Gates mit dem jeweiligen, prozentualen Anteil an Zellen.

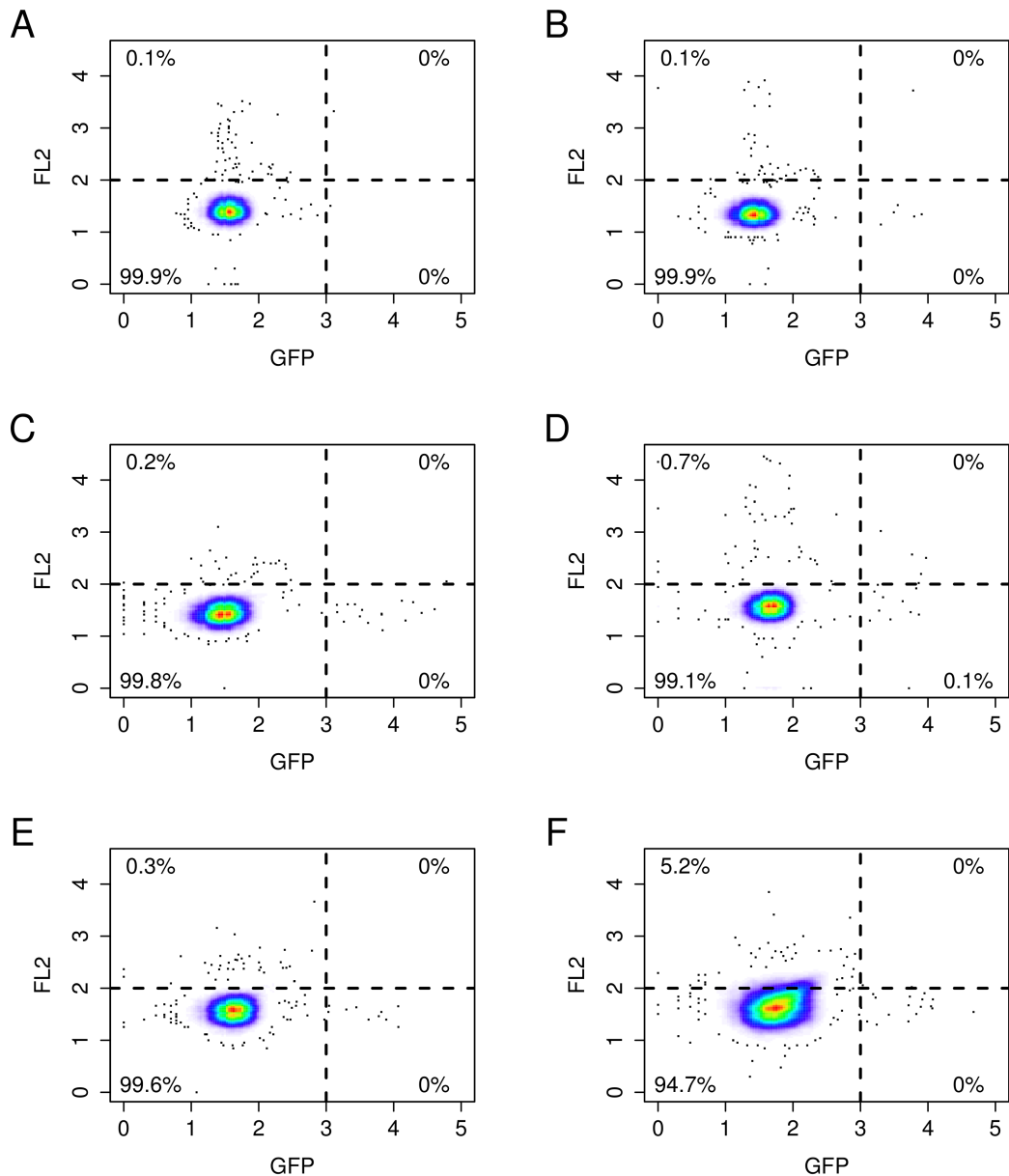


Abbildung 36: Dichteplots vom Wildtypstamm DSM 319. Die rote Fluoreszenz (FL2) wurde gegen die Fluoreszenz von eGFP im zeitlichen Verlauf nach der Induktion mit Xylose aufgetragen. A) Vor der Induktion und B) – F) 1 bis 5 Stunden nach der Induktion. Die gestrichelten Linien markierten die Auftrennung in vier Gates mit dem jeweiligen, prozentualen Anteil an Zellen. Im Verlauf einer Wachstumskurve konnte keine Heterogenität gemessen werden, lediglich in der stationären Phase erhöhte sich die Fluoreszenz etwas durch ein vermehrtes Leuchten der Septen (F).

Ein analoges Resultat zeigte auch das als Kontrolle dienende Leerplasmid pMM1520 (Abb. 37).

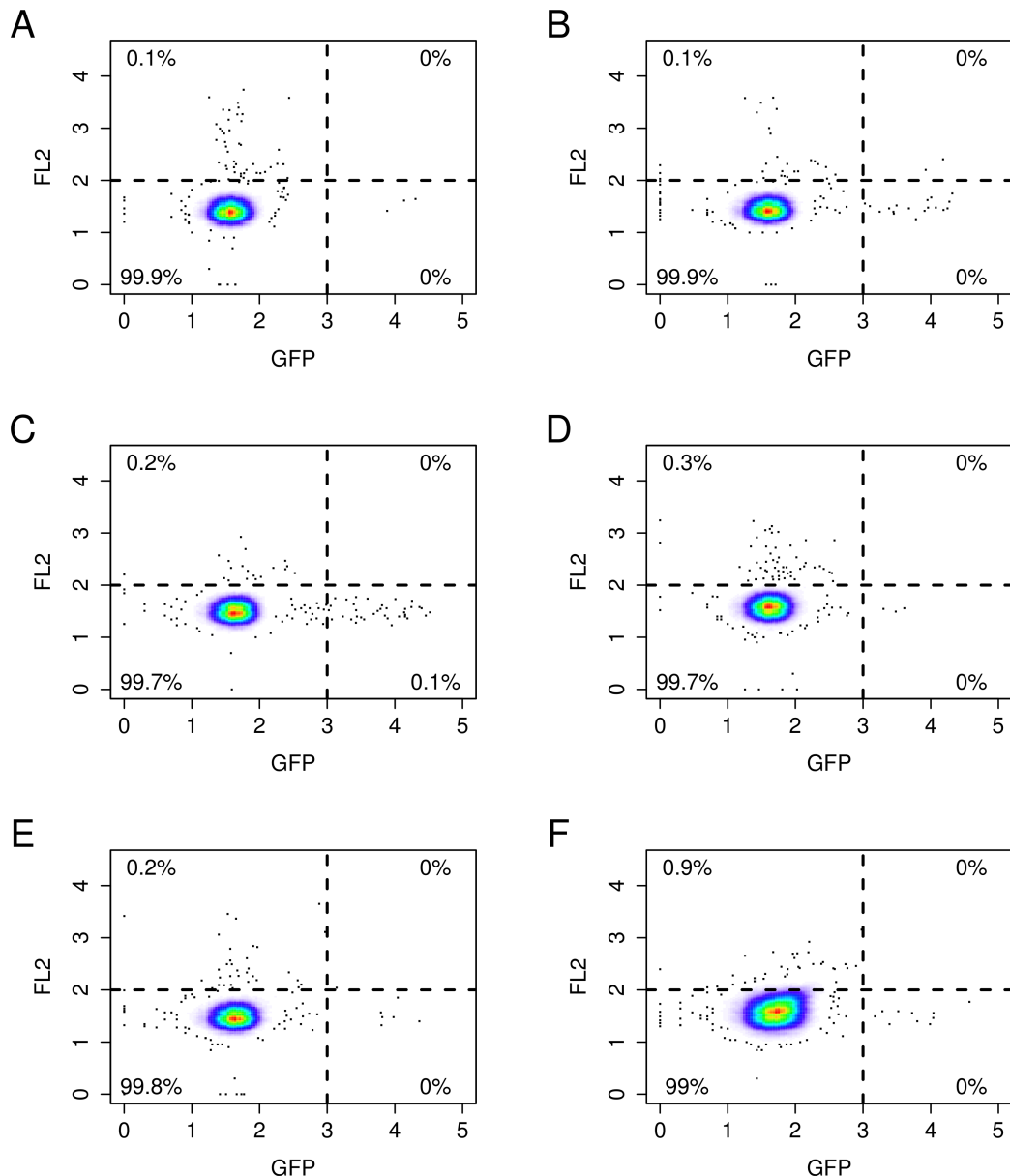


Abbildung 37: Durchflusszytometrische Daten des Kontrollstammes pMM1520, ein Derivat des Produktionsstammes, das neben den Resistenzen (Amp und Tet) *xyIR* enthält, aber nicht *P_{xyIA}-gfp*. Die Rotfluoreszenz (FL2) wurde gegen die Fluoreszenz von eGFP im zeitlichen Verlauf nach der Induktion mit Xylose aufgetragen. A) Vor der Induktion und B) – F) 1 bis 5 Stunden nach der Induktion. Die gestrichelten Linien markierten die Auftrennung in vier Gates mit dem jeweiligen, prozentualen Anteil an Zellen. Es konnte keine Heterogenität dieser Zellen festgestellt werden, die als Kontrolle für einen Stamm mit Leerplasmid dienen sollten.

In Zellen, die das Plasmid pKMMBm2 (Abb. 38) enthielten, konnte über das XylR-mCherry-Fusionsprotein ein zweiter Fluoreszenzkanal (FL2) gemessen werden. Das rote Fluoreszenzsignal war hier jedoch sehr schwach, so dass somit eine Trennung in einzelne Subpopulationen nicht ohne weiters möglich war. Die mit freundlicher Unterstützung

von Dr. Lothar Gröbe am HZI gewonnen FACS-Daten bestätigten dieses Resultat. Zum Sortieren konnten deshalb nur sichere Fraktionen im Randbereich verwendet werden. Auch dieser Stamm zeigte, wie der pSSBm85 Stamm, bereits vor der Induktion eine ba-

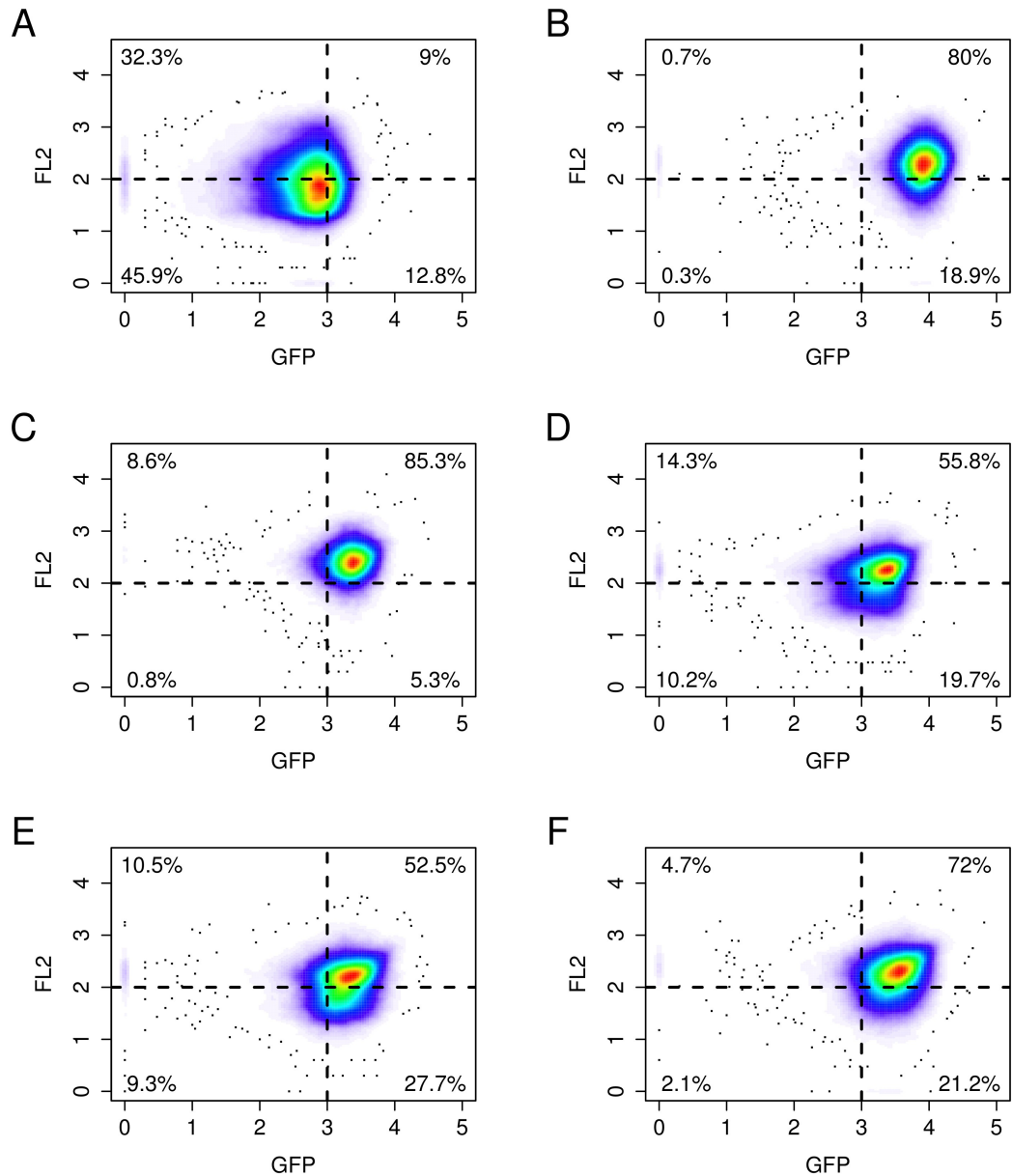


Abbildung 38: Dichteplots des Stamms pKMMBm2, der die XylR-mCherry-Fusion beinhaltet. FL2 wurde gegen die Fluoreszenz von eGFP im zeitlichen Verlauf nach Xyloseinduktion aufgetragen. A) Vor der Induktion und B) – F) 1 bis 5 Stunden nach der Induktion. Die gestrichelten Linien markierten die Auftrennung in vier Gates mit dem jeweiligen, prozentualen Anteil an Zellen.

sale Aktivität des Promotors (38A), so dass bereits nach 1 Std fast alle gemessenen

Ereignisse sowohl eGFP, als auch xylR-mCherry enthielten (38B). Im Folgenden kam es zu einer leichten Verschiebung auf die Seite des Signals von mCherry (38C und 38D), das anschließend wieder zurück wechselte (38E und 38F). Rein weiße Zellen waren nach Induktion somit kaum vorhanden.

Wie zu erwarten produzierte die XylR-Mutante pKMMBm5 bereits ohne Xylose eGFP (Abb. 39). Mit dem Beginn der Induktion bildeten sich zeitgleich verschiedene Subpopulationen aus. Zwischen der grünen und der weißen Fraktion stellte sich nach drei Stunden zunächst ein Gleichgewicht aus annähernd gleichen Anteilen von hoch- und niedrigproduzierenden Zellen ein (Abb. 39D). Gegen Ende der Wachstumskurve verschob sich dieses Gleichgewicht wieder auf die Seite der grünen Population (Abb. 39). Die größte Heterogenität und gleichzeitig das komplexeste Verhalten zeigte der Stamm mit den multiplen *xylR*-Bindestellen, pKMMBm6 (Abb. 40). Die Subpopulation, die mCherry enthielt war zunächst sehr gering. Vermutlich band der Repressor an die multiplen Bindestellen und verlor daher die Mobilität in der Zelle. Diese besaßen dann keinen schwach roten Hintergrund mehr, sondern erschienen weiß. Zudem könnte auch das Plasmid seine Mobilität durch die zusätzlich gebundenen Repressoren verlieren, so dass Zellen die wenig Plasmide besitzen am ehesten weiterwachsen und das Plasmid zwar replizieren, nicht aber für die Produktion garantieren. Evtl. ist es für die plasmidgefüllte Zelle mit gebundenem XylR schwieriger, die Produktion in Gang zu setzen. Ursprünglich sollte mit diesem Plasmid zum einen freies XylR wegtitriert werden und zum anderen eine verbesserte Visualisierung des Vektors über xylR-mCherry stattfinden. Der Anteil an rot fluoreszierenden Zellen ist in diesem Vektor dadurch deutlich geringer, als beim Stamm pKMMBm2, was in Abbildung 38D und 38E bestätigt wird. Die Subpopulation der weißen Zellen war im Gegenzug dazu in allen Wachstumsphasen extrem hoch. Am auffälligsten war, dass dieser Stamm ab Mitte der exponentiellen Phase gut abgrenzbare Subpopulationen aus weißen, roten, grünen und rot-grünen Zellen ausbildete. Möglicherweise führte hier der Effekt des Heraustitrierens von freiem XylR zu Zelllinien mit extremen Zahlenverhältnissen von XylR und Plasmidmolekülen. Darüber hinaus bereitete es große Probleme, von diesem Stamm Plasmid-DNA zu isolieren, da sich diese nicht vom Proteinkomplex trennen ließ und die Ausbeute dadurch extrem niedrig war. Um zu überprüfen, ob es sich bei der Extraktion um DNA handelte, wurde eine Retransformation in *E. coli* durchgeführt, was aber keinen Vorteil brachte. Zwar

konnten dabei etliche Klone generiert werden, die beschriebene Problematik bezüglich der Extraktion von DNA blieb jedoch weiterhin bestehen.

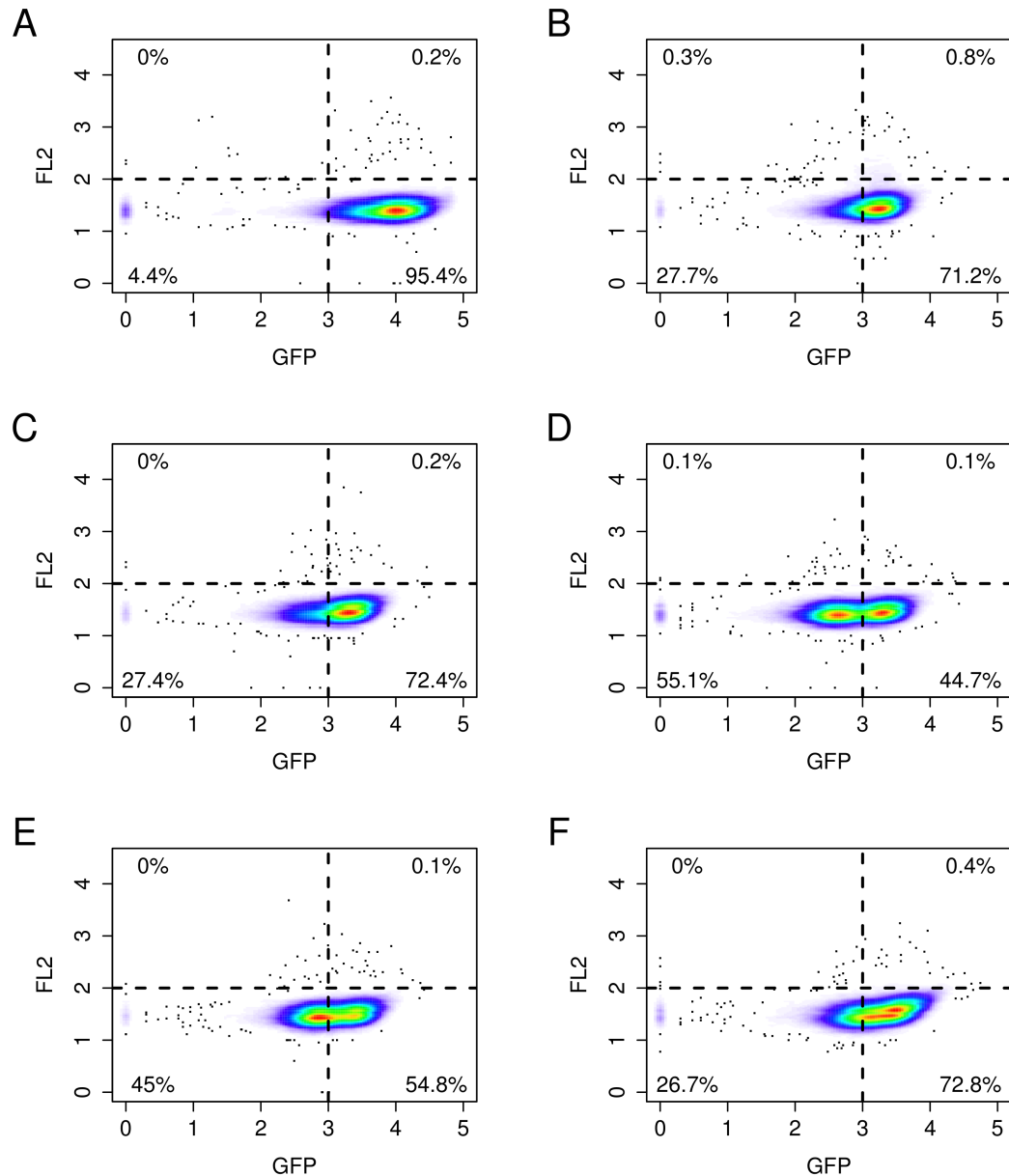


Abbildung 39: Durchflusszytometrische Ergebnisse der plasmidbezogenen *xyIR*-Mutante pKMMBm5. Der Fluoreszenzkanal im Rotbereich (FL2) wurde gegen die Fluoreszenz von eGFP im zeitlichen Verlauf nach Xyloseinduktion aufgetragen. A) Vor der Induktion und B) – F) 1 bis 5 Stunden nach der Induktion. Die gestrichelten Linien markierten die Auftrennung in vier Gates mit dem jeweiligen, prozentualen Anteil an Zellen.

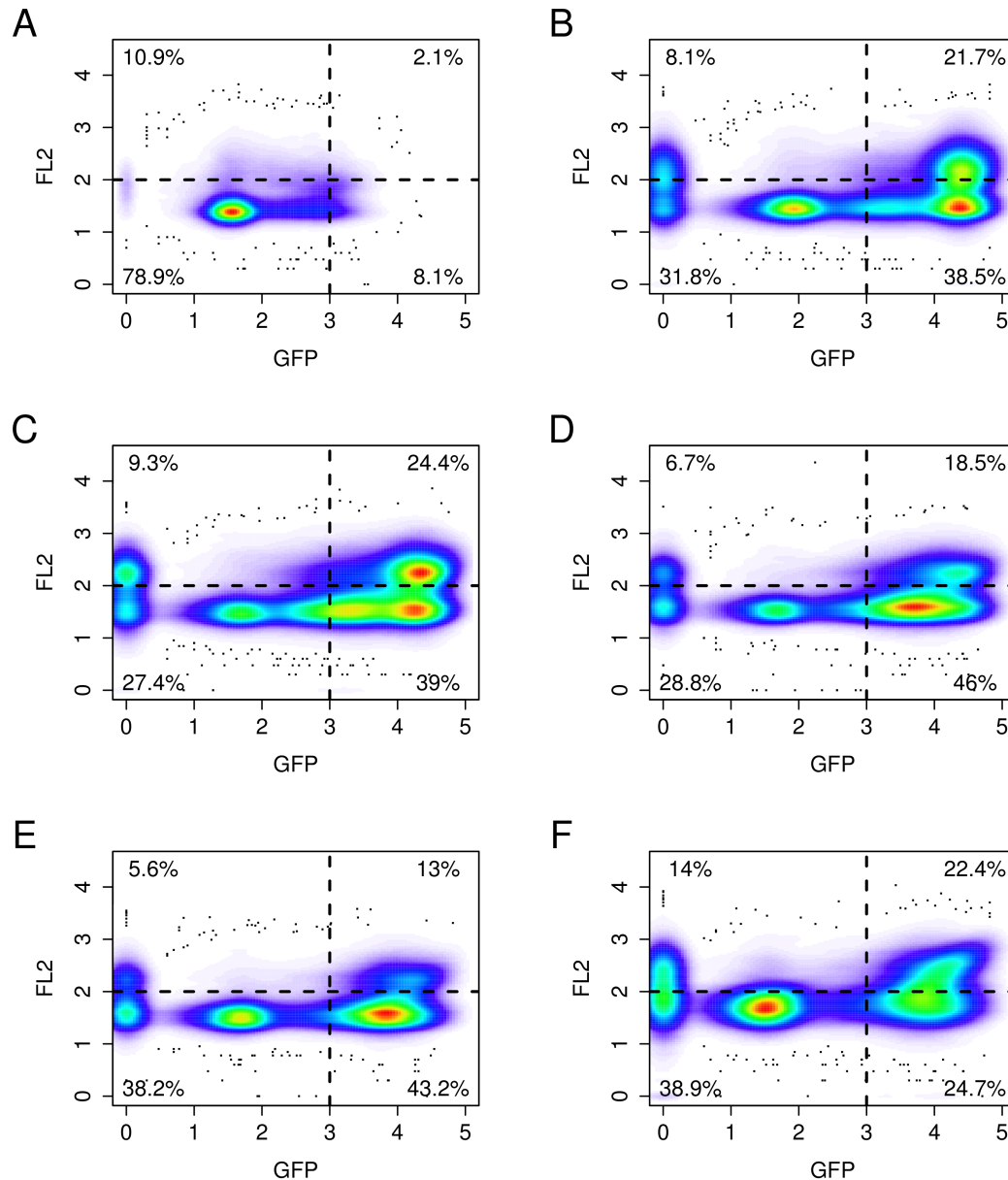


Abbildung 40: Dichtplots des Stamms pKMMBm6, dessen Plasmid multiple *xylR*-Bindestellen enthält. Die durchflusszytometrischen Daten zeigen die Fluoreszenz im Rotbereich (FL2) aufgetragen gegen das Signal von eGFP im zeitlichen Verlauf nach der Induktion mit Xylose. A) Vor der Induktion und B) – F) 1 bis 5 Stunden nach der Induktion. Die gestrichelten Linien markierten die Auftrennung bezüglich der Produktivität von eGFP und mCherry in vier Gates mit dem jeweiligen, prozentualen Anteil an Zellen.

3.6.2 Messung der Vitalität während der heterologen Proteinproduktion von eGFP in A5- und LB-Medium

In den gezeigten Zeitraffer-Mikroskopie-Experimenten führte eine ungleichmäßige Verteilung von Plasmiden zu einer sofortigen Verringerung der Produktionsraten und der Bildung von weißen, niedrigproduzierenden Tochterzellen. Aufgrund der Tatsache, dass die meisten Zellen in der stationären Phase überwiegend aus hochproduktiven Zellen bestehen, könnte dieses Ereignis theoretisch vernachlässigt werden. Allerdings besteht bei der Verwendung dieser Expressionssysteme im Chemostaten oder Fed-Batch-Bioreaktor ein Problem. Dort ist permanent eine stabile Subpopulation von Zellen vorhanden, die produktions schwach sind (Biedendieck *et al.*, 2007), was für industrielle Produktionsprozesse mit immensen finanziellen Einbußen verbunden sein kann. Darüber hinaus stagnieren plasmidgefüllte Zellen und stellen ihren Produktionsprozeß ein, bis sie schließlich beim Verwenden einer Batch Kultur sterben.

Erste Hinweise auf einen möglichen Zelltod dieser stagnierenden Zellen gab der DNA-Microarray (Kap. 3.9). Des Weiteren konnte insbesondere der Stamm, der das Plasmid pKMMBm5 trug, nur sehr schwer über längere Zeit vital gehalten werden. Hierfür war es unumgänglich, von einer frischen Platte und aus einer vitalen Vorkultur anzupfen. Um diese Beobachtungen aufzuklären, wurden vom Stamm DSM319 einschließlich aller verwendeten Plasmide Zellen aus einer wachsenden Kultur nach 5 Std, 14 Std und 24 Std auf xylosehaltiges A5-Medium ausplattiert. Der Wildtypstamm war dabei in seiner Vitalität kaum beeinträchtigt, wohingegen plasmidtragende Stämme zu jedem Zeitpunkt merklich weniger Wachstum zeigten. Nach 24 Std verschwanden die Kolonien ganz. Eine Ausnahme bildete pKMMBm5, dort gab es auch nach dieser Zeit noch ein minimales Wachstum, wobei ca. 98,5 % der Kolonien grün waren. Dieses Resultat deckte sich mit dem Zeitraffer-Film (Film 5 Anhang 8.1). Plattierte man diese Zellen abermals aus, blieb der Ursprungszustand erhalten, weiße Kolonien blieben weiß und grüne Kolonien grün. Es bildeten sich nicht die gewohnten Subpopulationen aus. Die grünen Zellen stagnierten – siehe Film – in der stationären Phase, um anschließend wieder zu wachsen. Betrachtete man diese grünen Kolonien unter dem Mikroskop, so zeigten die wachsenden Bakterienketten eine stark heterogene Population.

In einem weiteren Experiment, das im BioLector durchgeführt wurde, konnte sowohl

das Bakterienwachstum, als auch die eGFP-Produktion im halbdefinierten Minimalmedium A5 mit dem Vollmedium LB verglichen werden. Das Resultat ist in Abbildung 38 dargestellt. Bei Wachstum mit A5 nahm die absolute eGFP-Menge stetig ab, bis schließlich keine Fluoreszenz mehr gemessen werden konnte (Abb. 41A).

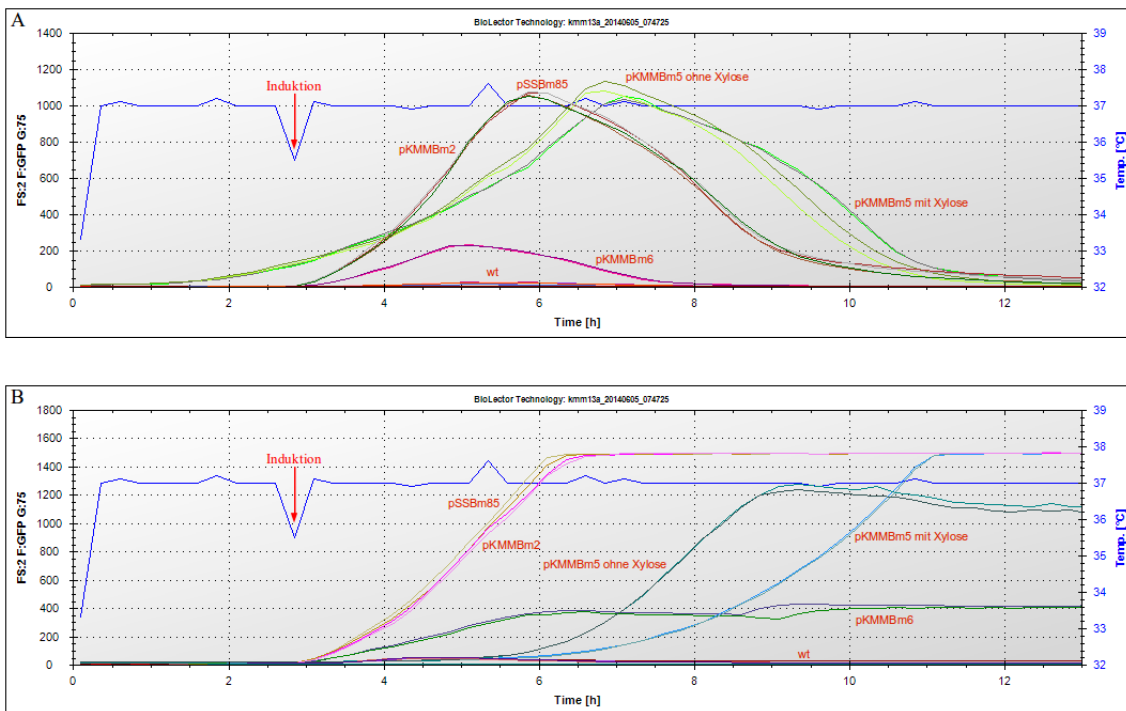


Abbildung 41: Stabilität von eGFP unter Verwendung diverser Medien. A) Bei Wachstum der plasmidtragenden Stämme im halbdefinierten Minimalmedium A5 nahm die Fluoreszenz und damit die Menge an eGFP ab, wohingegen B) im Komplexmedium LB die Produktion zwar eingestellt wurde, der Gehalt an eGFP aber mindestens 6 Stunden stabil blieb.

Interessanterweise setzte die Abnahme der Fluoreszenz mit dem Beginn einer scheinbar zweiten Wachstumsphase ein. Allerdings standen diese Werte entgegen der mit dem Photometer gemessenen optischen Dichte. Diese widerlegten das Auftreten einer zweiten Wachstumsphase und somit ein mögliches diauxisches Verhalten. Im BioLector wird die Biomasse über eine Streulichtmessung bestimmt. Möglicherweise wird das Streulicht durch die strukturelle Veränderung von Zellen beim Absterben beeinträchtigt und suggerieren eine zweite Wachstumskurve. Unter Verwendung von LB-Medium hingegen trat dieser scheinbar zweite Wachstumsschub weder im BioLector, noch mit der photometrisch bestimmten optischen Dichte auf. Hinsichtlich des Gehalts an eGFP

stellte sich beim Inokulieren in LB-Medium ein Maximum ein, das über mehrere Stunden stabil blieb (Abb. 41B). Dabei konnte das Auftreten von Photobleaching ausgeschlossen werden, da eGFP in LB-Medium weiterhin stabil blieb. Die Stabilität der Fluoreszenzintensität von plasmidtragenden Stämmen, die zudem mCherry enthielten blieb im Gegensatz dazu bei beiden Medien konstant und nahm lediglich in A5-Medium mit Eintritt in die stationäre Phase minimal ab, hielt sich dann aber konstant auf diesem Level.

Die Mikrotiterplatten des BioLectors wurden nach dem Lauf makroskopisch verglichen. Zunächst unterschieden sich die beiden unterschiedlichen Nährmedien beim Betrachten im normalen Licht bezüglich ihrer Farbe (Abb. 42A). Während die Kulturen in A5 eine milchig weiße Trübung aufwiesen (Abb. 42A linke Seite), zeigten die in LB gewachsenen Bakterien eine trüblich grüne Farbe (Abb. 42A rechte Seite), wobei die induzierten (Doppelbestimmung rechts außen) und die nichtinduzierten Kulturen, die das Plasmid pKMMBm5 trugen eine tiefgrüne Färbung aufzeigten. Das Ergebnis mit einer Anregung

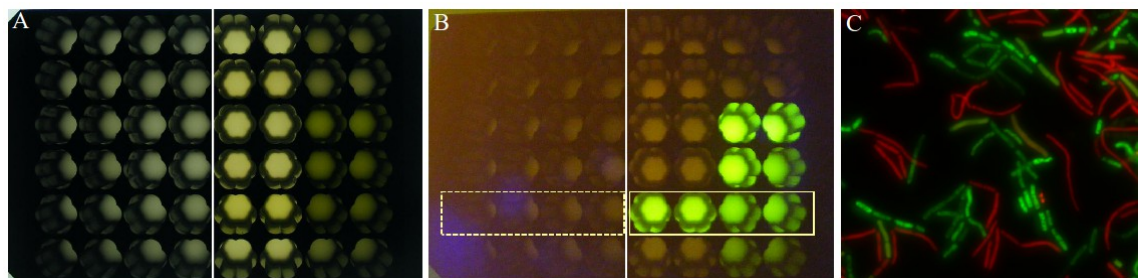


Abbildung 42: Vitalitätsverlust der verwendeten Kulturen. A) und B) Von oben nach unter wurde folgende Probenreihenfolge eingehalten: DSM 319, pMM1520 (Leerplasmid), pSSBm85 (Produktionsstamm), pKMMBm2 (XylR-mCherry-Fusion), pKMMBm5 (*xylR*-Mutante bezüglich des Plasmids) und pKMMBm6 (multiple *xylR*-Bindestellen). Im jeweils linken Bereich fand die Inkubation mit halbdefiniertem A5-Medium statt, rechts hingegen wurde mit LB-Medium (Komplexmedium) inokuliert. Die beiden ersten Doppelbestimmungen zeigen jeweils die Kultur ohne Induktion, die beiden letzten jeweils nach Zugabe von Xylose. A) Zunächst wurde mit Weißlicht stimuliert, B) dann mit Blaulicht. Lediglich in LB-Medium gewachsene Bakterien fluoreszierten noch. C) Von dem rechten Teil der eingerahmten Probe pKMMBm5 wurde eine Lebend/Tot-Färbung mit Propidiumiodid durchgeführt. Bei toten bzw. lädierten Zellen dringt der Farbstoff in die Zelle ein und leuchtet rot. Die beiden umrandeten Proben wurden durchflusszytometrisch analysiert (Abb. 43).

im Blaulichtbereich (488 nm) geht aus Abbildung 42B hervor. Die induzierten Kulturen, die im halbdefinierten Medium wuchsen (linke Seite), zeigten keine eGFP-Fluoreszenz mehr, wohingegen die im Komplexmedium inkubierten Zellen weiterhin fluoreszierten,

sofern sie ein entsprechendes Gen für eGFP besaßen. Lediglich die plasmidbezogene *xyIR*-Mutante produzierte vor Induktion bereits eGFP (eingerahmte rechte Proben Abb. 42B). Von den beiden eingerahmten Proben wurde eine Lebend/Tot-Färbung mit Propidiumiodid (PI) durchgeführt und mikroskopiert. Lediglich die Bakterien in LB-Medium enthielten noch vitale und grüne Zellen (Abb. 42C). Diese, die in A5-Medium wuchsen waren komplett tot. Die eingerahmten Proben mit dem Plasmid pKMMBm5 wurden zusätzlich mit Hilfe des Durchflusszytometers analysiert (Abb. 43).

Vitale, eGFP-produzierende Zellen können den roten Farbstoff Propidiumiodid aufgrund der intakten Zellmembran nicht aufnehmen (Abb. 43A). Komplette tote Zellen, die aus Abbildung 42B (eingerahmte Proben links im Bild) stammten waren überwiegend rot-fluoreszierend (Abb. 43B). Noch grün leuchtende Zellen (eingerahmte Proben rechts im Bild in Abb. 42B) waren teilweise vital, zum Teil aber bereits schon tot (Abb. 43C). Diese Zellen könnten durch die immens hohe Produktionsleistung gestresst worden sein, bis sie schließlich starben. Extremer spitzte sich diese Situation offensichtlich noch bei der Verwendung von Minimalmedium zu, was auf einen Hungerzustand der Zellen hindeutet (siehe auch Ergebnis des DNA-Microarrays) und eGFP entweder ab- oder umgebaut wird.

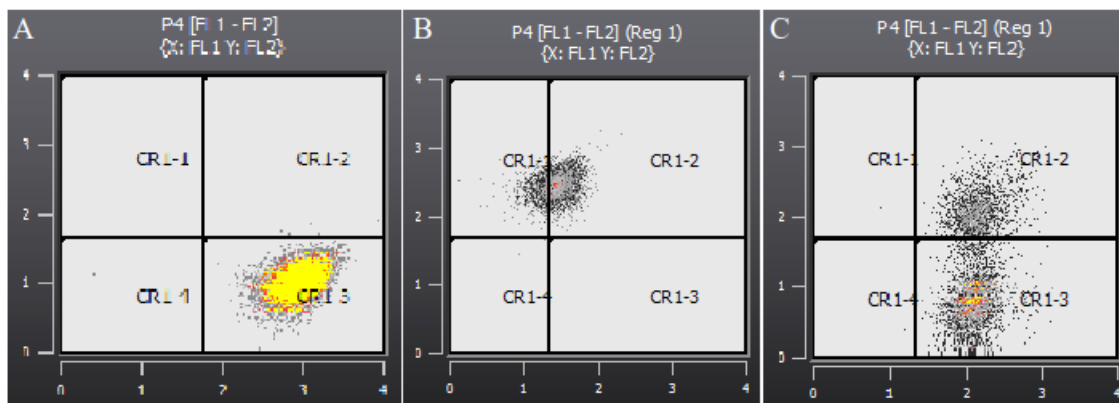


Abbildung 43: Lebend/Tot-Färbung des Stammes DSM 319, der das Plasmid pKMMBm5 trägt. Die x-Achse beschreibt den grünen Fluoreszenzkanal (FL1) für eGFP und die y-Achse den roten Fluoreszenzkanal (FL2) für mCherry. A) Zunächst ist eine vitale eGFP-produzierende Kultur zu sehen (rechtes unteres Quadrat. B) Komplette tote Zellen befinden sich im oberen Bereich (Quadranten CR1-1 und CR1-2). Diese stammten von der eingerahmten, nicht fluoreszierenden Probe aus Abbildung 41B. C) Eine noch grün fluoreszierende Probe (Abb. 41B rechte Einrahmung) beinhaltete sowohl lebende, als auch tote Zellen (siehe auch 39C).

Im Folgenden sollen deshalb die Situation einer wachsenden Batch Kultur des Wildtypstammes und seiner plasmidtragenden Stämme genauer mit dem Durchflusszytometer analysiert werden. Zunächst wurde eine PI-Färbung des Wildtypstammes durchgeführt (Abb. 44).

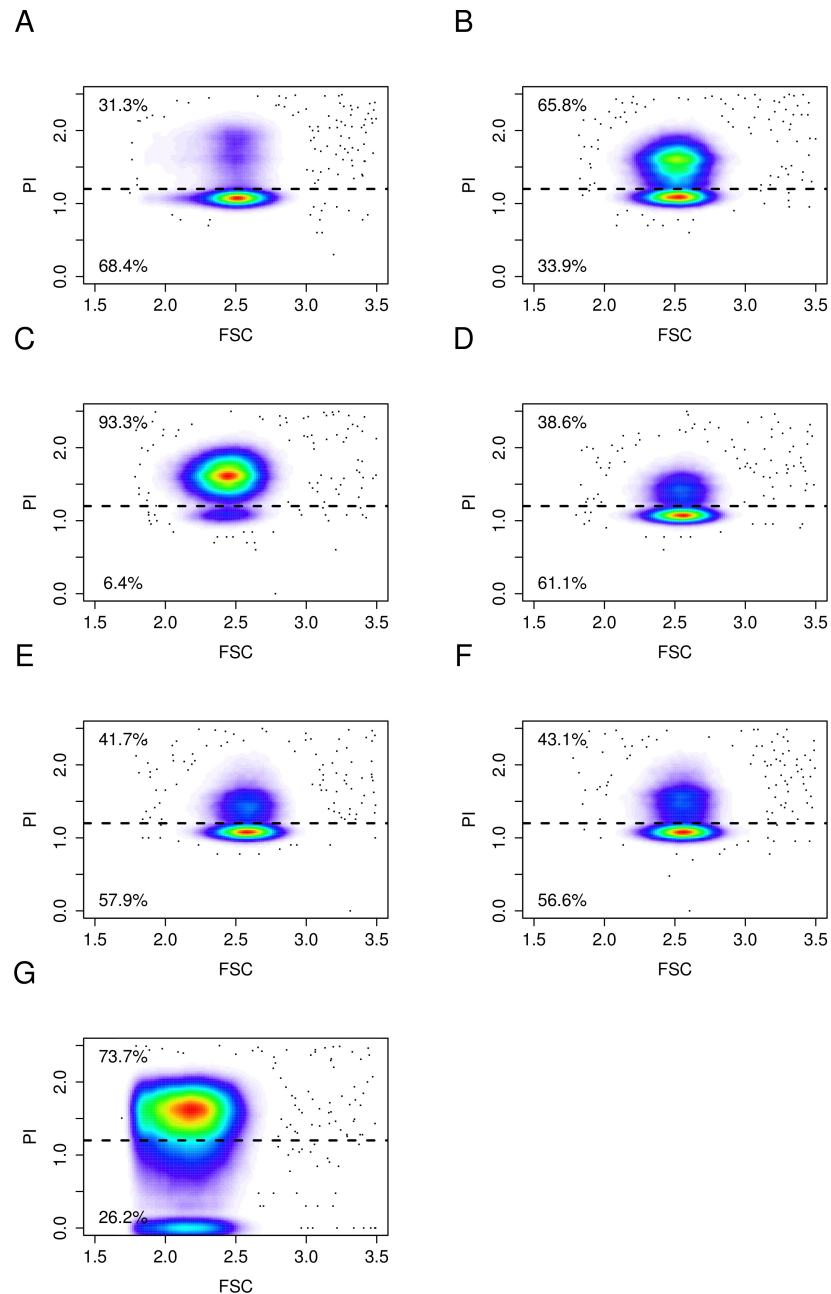


Abbildung 44 Lebend/Tot-Färbung mit PI des Wildtypstammes DSM 319. Die durchflusszytometrischen Daten zeigen FSC aufgetragen gegen PI im zeitlichen Verlauf nach der Induktion mit Xylose. A) Vor der Induktion und B) – G) 1 bis 6 Stunden nach der Induktion. Die gestrichelte Linie markierte die Auftrennung bezüglich der vitalen und toten Subpopulationen in zwei Gates mit dem jeweiligen, prozentualen Anteil an Zellen.

Jedoch war bereits zu Beginn einer Kultur mit ca. 31 % ihre Vitalität eingeschränkt (Abb. 44A), was vermutlich an der vorausgehenden Vorkultur lag. Während der exponentiellen Phase stieg der Anteil an roten Zellen erheblich an und lag bei ca. 93 % (Abb. 44B und 44C). Es etablierten sich zwei Maximas zwischen den vitalen und den bereits lädierten Zellen. Im weiteren Verlauf der Meßreihe sank dieser Anteil zunächst wieder auf ca. 40 % drastisch ab (Abb. 44D und 44E), um dann anschließend wieder anzusteigen. Am Ende wurde ein Zustand erreicht, bei dem ca. 26 % der Zellen aus agilem Zellmaterial bestand sowie 74 % aus solchen mit geringerer Vitalität (Abb. 44F).

Im nächsten Schritt wurde der Einfluss von Tetracyclin auf die Vitalität untersucht. Betrachtete man die PI-Färbung des Produktionsstammes pSSBm85 (Abb. 45), erzeugten bereits vor Xylosezugabe mehr Zellen eGFP in Medium mit Tetracyclin, als die in Abbildung 33A vorhandene basale Produktion (Abb. 45A). Im weiteren Verlauf nahm die Menge an produktiven Zellen zu (Abb. 45B). Nach einer Abnahme der eGFP-Menge (Abb. 45 C) produzierten im Folgenden mehr als die Hälfte aller Zellen (Abb. 45D und 45E). Ein Höhepunkt konnte schließlich 5 Std nach der Induktion erreicht werden (Abb. 45F). Nach 6 Std wurde abermals ein Zustand vorgefunden, bei dem kein eGFP mehr detektierbar war (Abb. 44G). Am Anfang der frühen exponentiellen Phase war die Vitalität erneut eingeschränkt (Abb. 44A), was vermutlich auch hier wieder an der vorausgehenden Vorkultur im Minimalmedium A5 lag. Während der exponentiellen Phase stieg der Anteil an roten PI-gefärbten Zellen erheblich an (Abb. 45B und 45C). Es etablierten sich zwei Maximas zwischen vitalen und lädierten Zellen (Abb. 45C), wobei sich dann das Maximum auf die Seite der vitalen Zellen verschob (Abb. 45D) und mit einem schleichenden Abfall (Abb. 45E) sowie einer erneuten Zunahme an fitten Zellen (Abb. 45F) dort auch blieb. Innerhalb der stationären Phase konnte rund 41 % Zellmaterial gefunden werden, das in seiner Fitness beeinträchtigt war. Eine Zunahme der Produktion korreliert scheinbar mit einer abnehmenden Vitalität.

Begutachtete man das Verhalten des Produktionsstammes ohne Antibiotikum, so ergab sich ein ähnliches Muster mit anderen Zahlenverhältnissen (Abb. 46). Bereits 1 Std nach der Zugabe von Xylose war eine immens hohe Anzahl der Zellen (etwa 75 %) nicht mehr vital (Abb. 46B). Dies steht im drastischen Gegensatz zu der Kultur mit Tetracyclin (Abb. 45). Offenbar wirkte sich das Antibiotikum ähnlich wie bei der Hormese vitalisierend auf die Zellen aus.

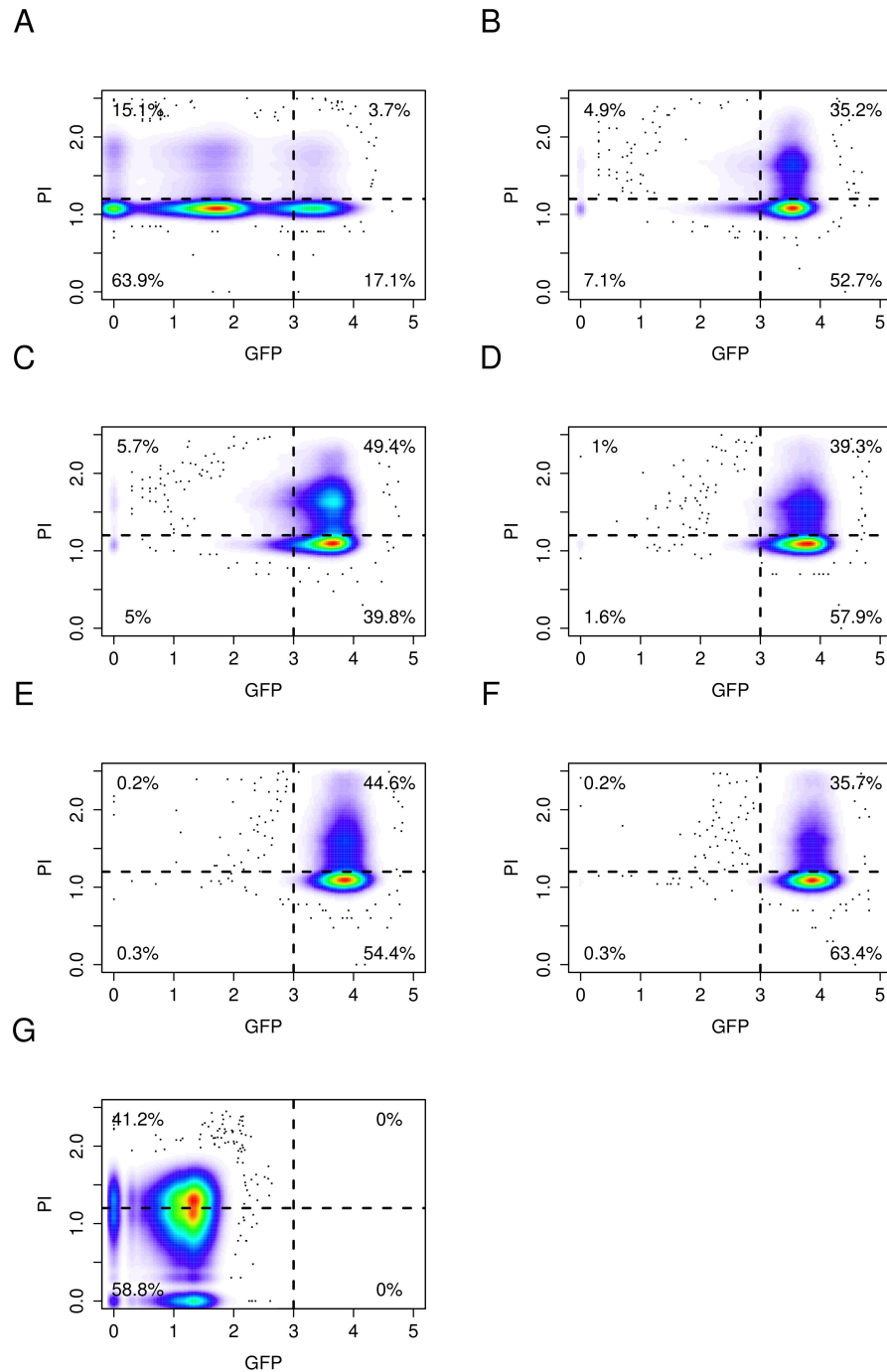


Abbildung 45: Durchflusszytometrische Daten einer PI-Färbung mit Tetracyclin des pSSBm85-Stamms. Die Dichteplots zeigen PI aufgetragen gegen eGFP im zeitlichen Verlauf nach der Induktion mit Xylose. A) Vor der Induktion und B) – G) 1 bis 6 Stunden nach der Induktion. Die gestrichelten Linien markierten die Auftrennung bezüglich der Produktivität von eGFP und Vitalität durch PI in vier Gates mit dem jeweiligen, prozentualen Anteil an Zellen.

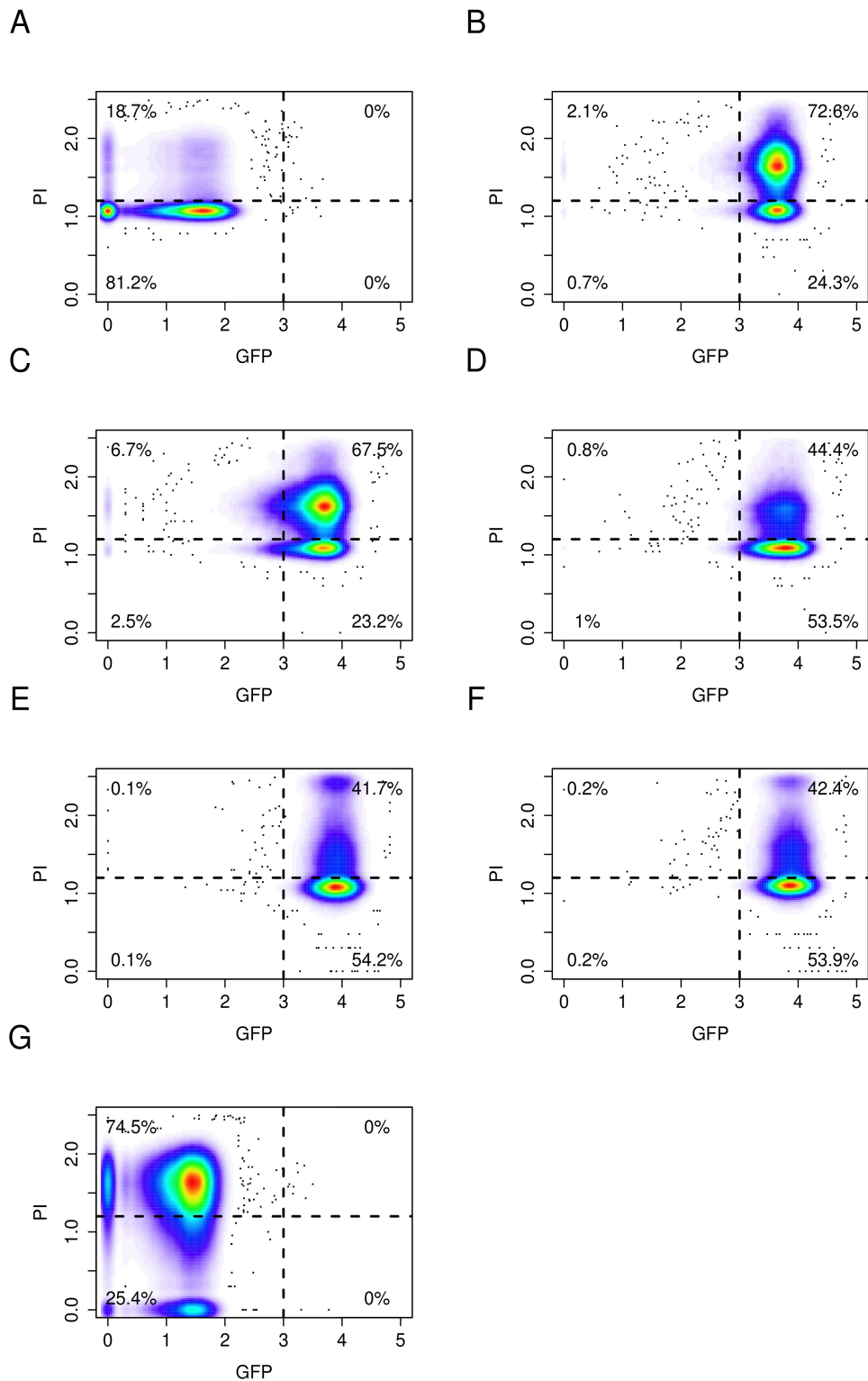


Abbildung 46: Durchflusszytometrische Daten der PI-Färbung ohne Tetracyclin vom Produktionsstamm pSSBm85. Die Dichteplots zeigen PI aufgetragen gegen eGFP im zeitlichen Verlauf nach der Induktion mit Xylose. A) Vor der Induktion und B) – G) 1 bis 6 Stunden nach der Induktion. Die gestrichelten Linien markierten die Auftrennung bezüglich der Produktivität von eGFP und Vitalität durch PI in vier Gates mit dem jeweiligen, prozentualen Anteil an Zellen.

Während die grüne Subpopulation einen Großteil der verfügbaren Energie für die Produktion aufwenden muss, könnte die weiße Population ihre Energie zunächst für die Replikation des Vektors verwenden, weiterwachsen und somit das Überleben sichern, indem Nachkommen gebildet werden, die abermals zwei Subpopulationen ausbildet. Darüber hinaus könnte eine wachsende Zelle stochastische Effekte kompensieren und das vorhandene Antibiotikum ausdünnen. Anfänglich erkannte man hier nur die basale eGFP-Produktion einiger Zellen und annähernd 19 % des Zellmaterials zeigte eine eingeschränkte Vitalität (46A). Im Anschluss konnte eine stetig zunehmende Rate an eGFP nach Induktion beobachtet werden (46B bis 46F). Mit dem Eintritt in die stationäre Phase konnte kein Produkt mehr detektiert werden (46G). Schon 1 Std nach der Induktion kam es mit 75 % zu einem drastischen Ansteigen an Bakterienzellen, die in ihrer Vitalität eingeschränkt waren (46B). Nach 2 Std verringerte sich lediglich der Anteil an Produzenten, die Fitness blieb gleich (46C). Es erfolgte eine Zunahme vitaler Zellen (46D und 46E) bis auf 54 % (46F). Innerhalb der stationären Phase waren ca. 75 % der Zellen tot (46G).

Ebenso bestätigte die Lebend/Tot-Färbung des Wildtyps DSM 319 (Abb. 44) das vorhergehende Resultat. Augenscheinlich wirkte sich dieser fehlende Selektionsdruck auf die Vitalität der Zellen aus und steigerte die mit PI angefärbte Fraktion. Zudem bildete sich kaum noch eine weiße Subpopulation aus. Den Färbungen, bei denen PI angewendet wurde, ist die fehlende ausgeprägte Heterogenität gemeinsam. Eine Möglichkeit wäre eine PI-bedingte Verschiebung der Heterogenität, da die PI-Aufnahme von der Wachstumsphase abhängig ist (Shi *et al.*, 2007).

Auch mit dem Leerplasmid pMM1520 wurde eine PI-Färbung durchgeführt (Abb. 47). Dieses Plasmid beinhaltet weder ein Gen für eGFP, noch für mCherry. Neben den Resistenzgenen enthält er nur das *xylR*-Gen. Dieser Stamm besaß durchgängig eine hohe Anzahl an Zellen geringer Vitalität. Auch hier stieg der Anteil bis auf 96 % (Abb. 47C) und fiel dann wieder auf ca. 79 % ab (Abb. 47F).

Die *xylR*-Mutante pKMMBm5 besaß zunächst in etwa 49 % an unfitem Zellmaterial, wobei davon aufgrund des fehlenden Repressors die größte Fraktion grün fluoreszierte (Abb. 48A). Mit der Zugabe von Xylose sank die Subpopulation der Niedrigproduzierer und die Bakterienkultur zeigte eine gesteigerte Fitness (Abb. 48B und 48C). Nach und

nach bildeten sich zwei (48B) bzw. vier Maximas (48C bis 48E) aus, die erneut auf eine Verschiebung der Heterogenität mittels PI hindeuten. Würde man die eGFP-Begrenzung des gesetzten Gates hypothetisch betrachtet von 3,0 auf 4,0 verschieben, so würde man vier verschiedene Subpopulationen erhalten, die eine vorkommende weiße Fraktion beinhalten würde. Am Ende der Kultivierung in der stationären Phase waren ca. 62 % der grünen Zellen tot (48F).

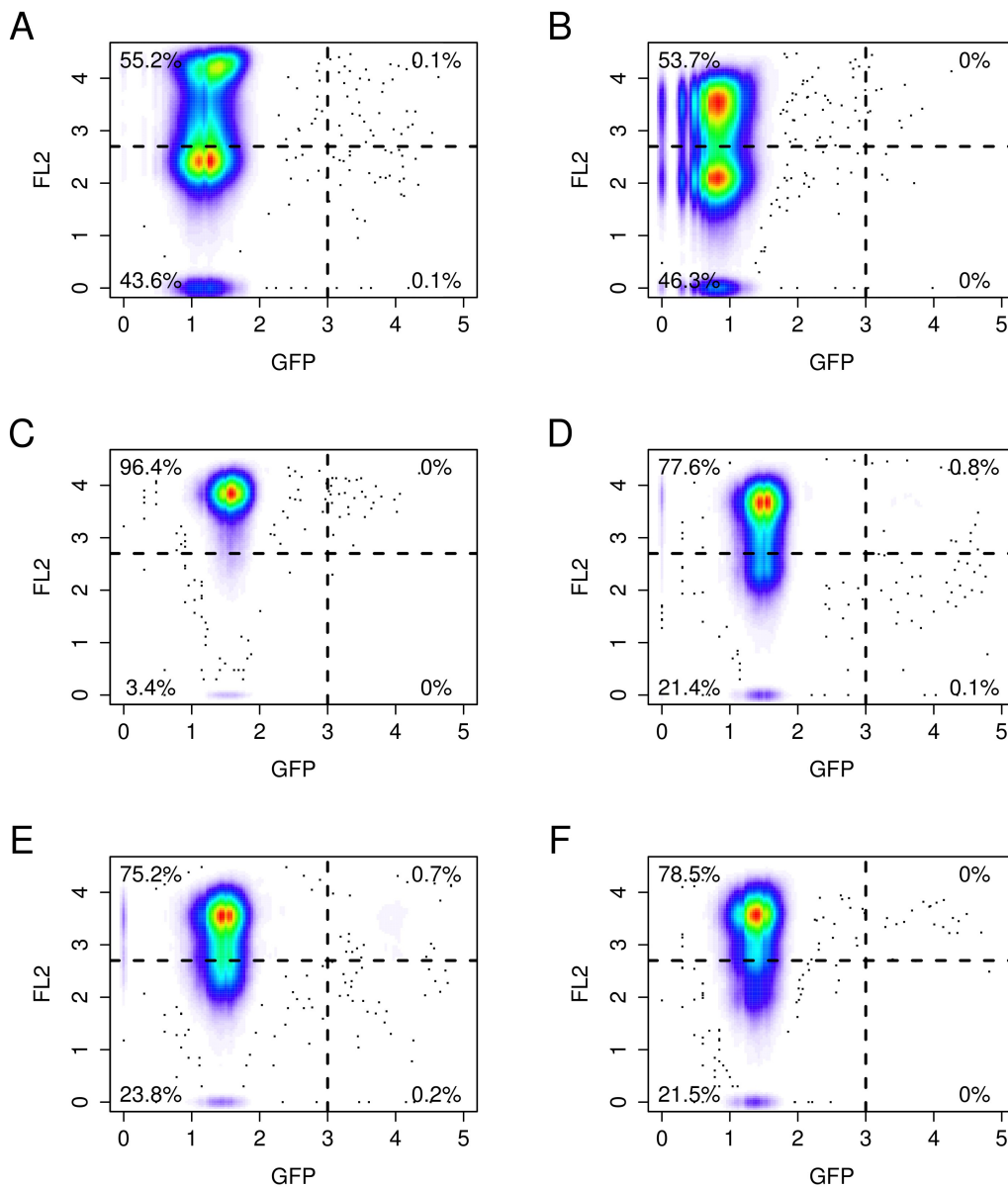


Abbildung 47: Durchflusszytometrie der PI-Färbung des Stamms pMM1520 (Leerplasmid). Die Dichteplots zeigen PI aufgetragen gegen eGFP im zeitlichen Verlauf nach der Induktion mit Xylose. A) Vor der Induktion und B) – F) 1 bis 5 Stunden nach der Induktion. Die gestrichelten Linien markierten die Auftrennung bezüglich der Produktivität von eGFP und Vitalität durch PI in vier Gates mit dem jeweiligen, prozentualen Anteil an Zellen.

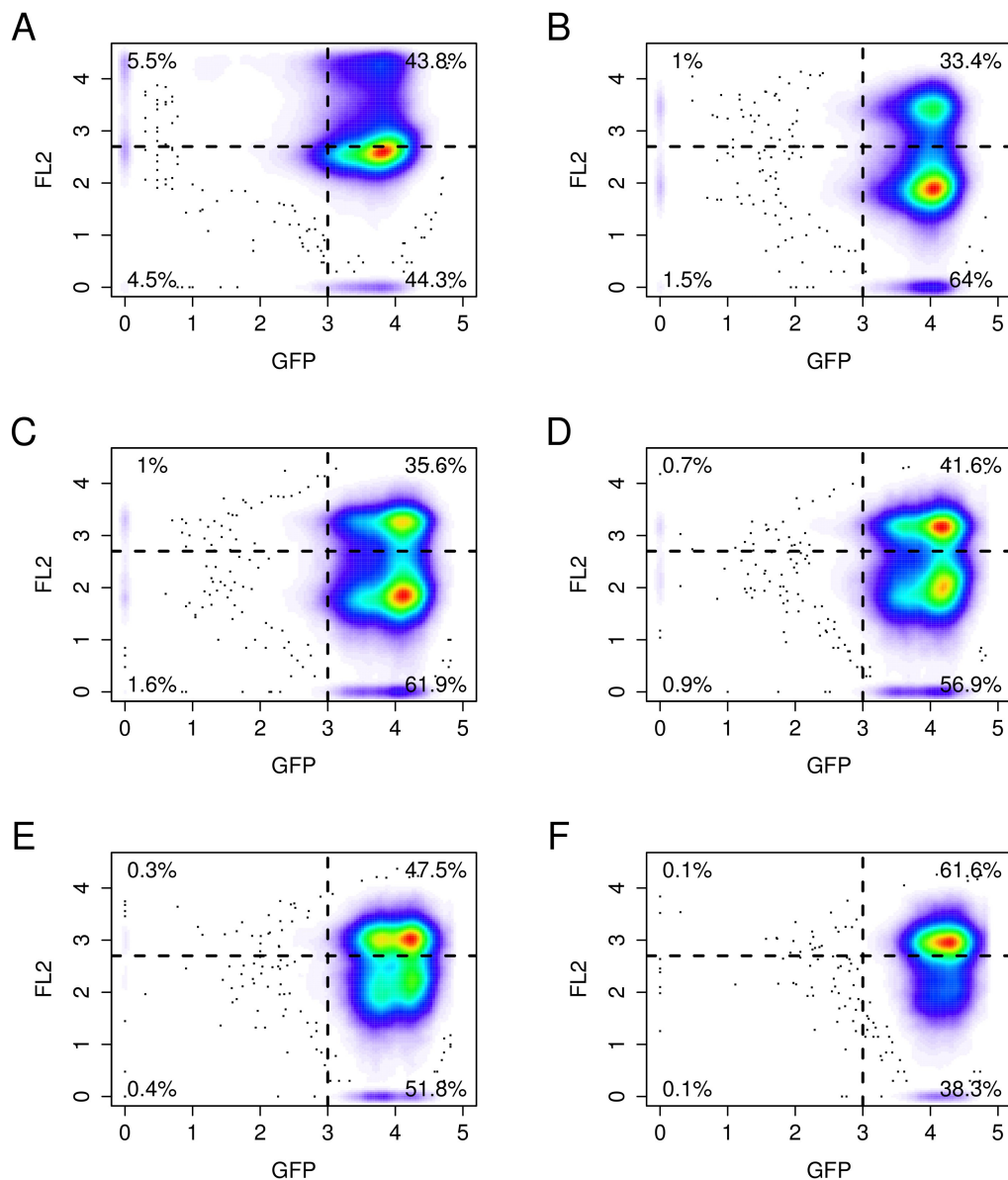


Abbildung 48: Durchflusszytometrische Daten der PI-Färbung des pKMMBm5-Stamms. Die Dichteplots zeigen PI aufgetragen gegen eGFP im zeitlichen Verlauf nach der Induktion mit Xylose. A) Vor der Induktion und B) – F) 1 bis 5 Stunden nach der Induktion. Die gestrichelten Linien markierten die Auftrennung bezüglich der Produktivität von eGFP und Vitalität durch PI in vier Gates mit dem jeweiligen, prozentualen Anteil an Zellen.

Insgesamt betrachtet nahm die Vitalität im Laufe einer Kultur nicht beständig ab. Sie stieg sogar innerhalb einer bestimmten Wachstumsphase wieder an (Abb. 43 bis 47). Zudem war auffallend, dass die Kulturen ohne Antibiotikum (DSM 319 und pSSBM85) sehr schnell einen hohen Anteil PI-gefärbter Messungen besaß, die auf lädierte Zellen hindeuteten (Abb. 43 und Abb. 45). Tetracyclin erhöht offenbar die Vitalität der ohnehin geschwächten Zellen im Minimalmedium.

3.6.3 Einfluss der Plasmidverteilung auf die Proteinproduktionsdynamik

Der mögliche Einfluss der Populationsheterogenität auf die heterologen Proteinproduktionsprozesse wurde systematisch analysiert. Zwei markante Effekte mit gewichtigen Einflüssen konnten beobachtet werden. Zum einen zeigten die Daten der Durchflusszytometrie, dass die Heterogenität während der ersten drei Stunden nach Xyloseinduktion am größten war. Mit Induktionsbeginn bestand die Subpopulation der leistungsstarken Produzenten zum einen aus Zellen mit einer hohen Kopienzahl des Plasmids, die im Laufe der Zeit immer mehr anstieg. Zum anderen wurde eine Subpopulation an Zellen gefunden, die extrem viel Plasmid ansammelte, bis dies schließlich zum Ende der Zellteilung führte. Zeitraffer-Filme verdeutlichten den Zusammenhang zwischen Zellen, die bereits früh mit der Produktion begannen, damit ein verringertes Wachstum zeigten und mitunter sogar aufhörten zu wachsen, während die weiße Subpopulation weiterwuchs und somit eine hemmende Wirkung auf den Produktionsprozeß hatte (Film 1 bis Film 5 Anhang 8.1).

Zur weiteren Untersuchung dieser Zusammenhänge wurde das Wachstumsverhalten und die Produktion an eGFP im Mikrobioreaktor (BioLector) konkretisiert (Abb. 49). Sämtliche biologische und technische Replikate hatten etwa 4 bis 5 Std nach der Induktion eine Stagnierung der Produktion von eGFP im halbdefinierten A5-Minimalmedium gemeinsam, wobei obendrein eGFP abgebaut wurde (Abb. 49B). Um eine schnelle Degradation oder Photobleich-Effekte ausschließen zu können, wurde die *in vivo* Halbwertszeit von eGFP bestimmt. Hierzu wurde die Proteinbiosynthese 1 Std nach der Induktion durch die Zugabe von Chloramphenicol abgestoppt und der eGFP-Abbau gemessen (Abb. 50). Mittels linearer Regression wurde eine minimale Halbwertszeit von 9,5 Std eruiert (Münch *et al.*, 2015). Unter Produktionsbedingungen scheint demnach ein aktiver Abbau von eGFP zu erfolgen.

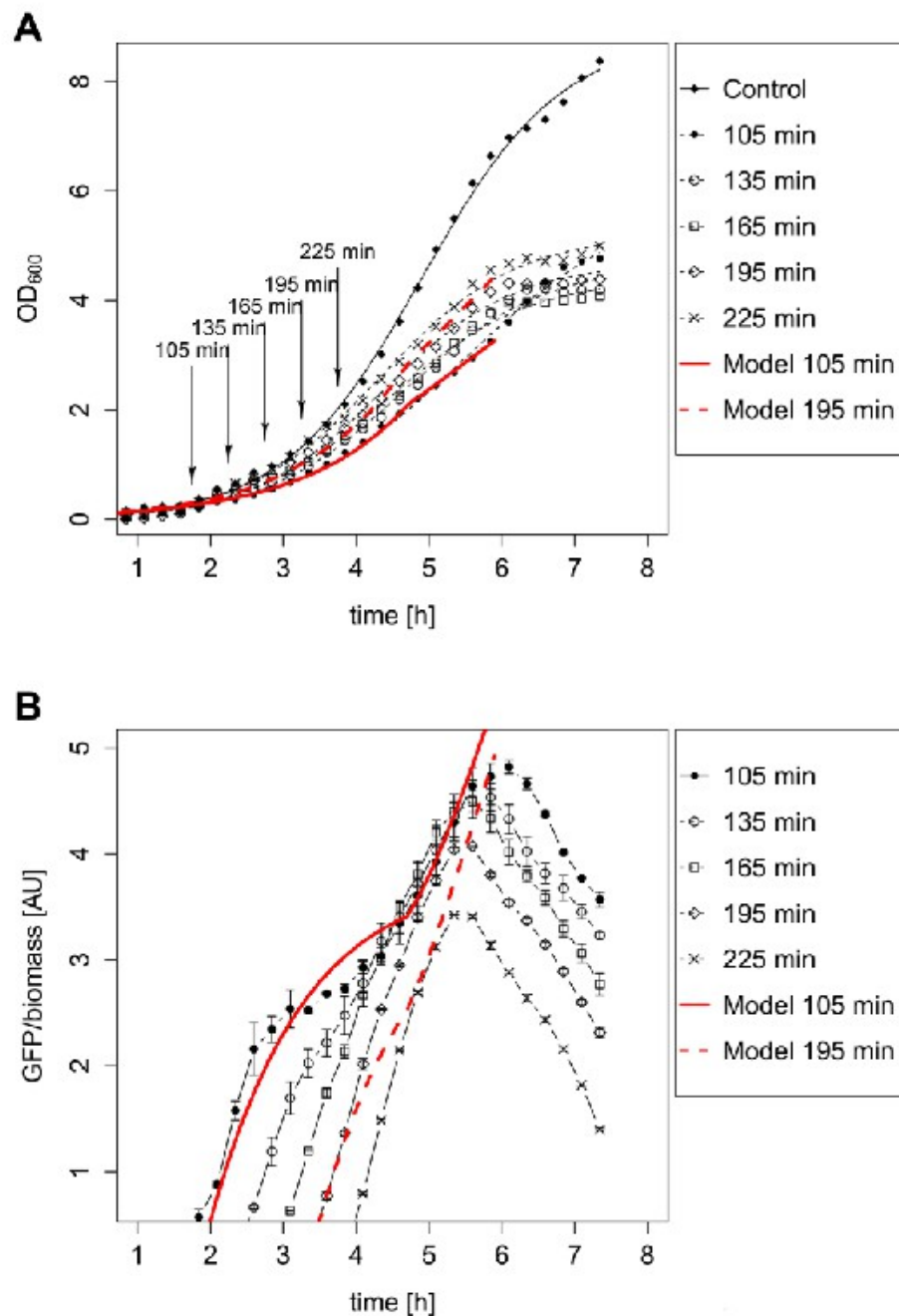
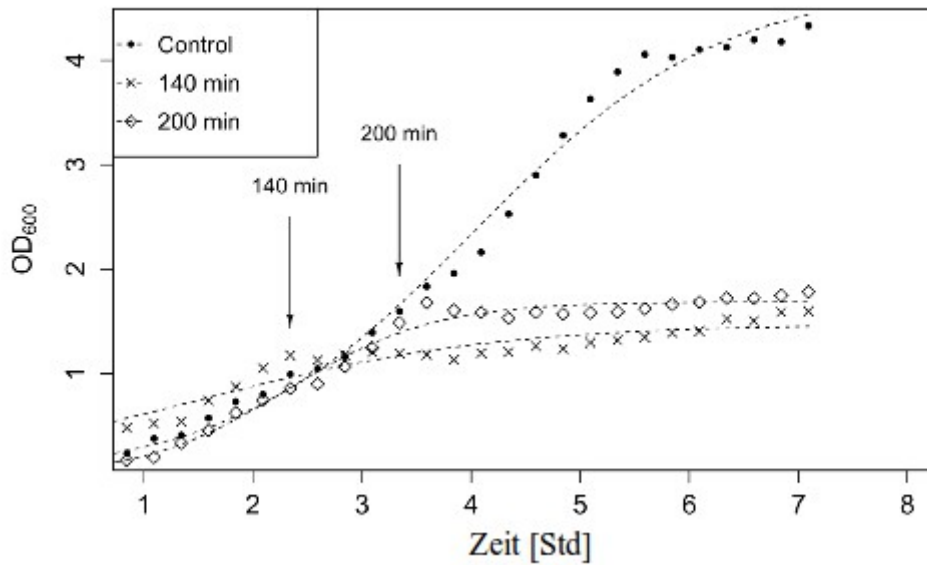


Abbildung 49: Die Produktion an eGFP vom Stamm pSSBm85 zu verschiedenen Induktionszeitpunkten. Die Kontrolle zeigt das Wachstum ohne Xylose und die übrigen schwarzen Kurven das Verhalten nach Induktion. A) Je später diese stattfand, desto mehr Biomasse wurde hergestellt. B) Nach der Normalisierung der eGFP-Produktion auf die Biomasse, konnte eine höhere Ausbeute bei früherer Induktion festgestellt werden. Die roten Kurven zeigen A) das modellierte Wachstum (Münch *et al.*, 2015) und B) das Produktionsverhalten zu einem Induktionszeitpunkt von 105 min (durchgezogene Linie) und von 195 min (gestrichelte Linie).

A



B

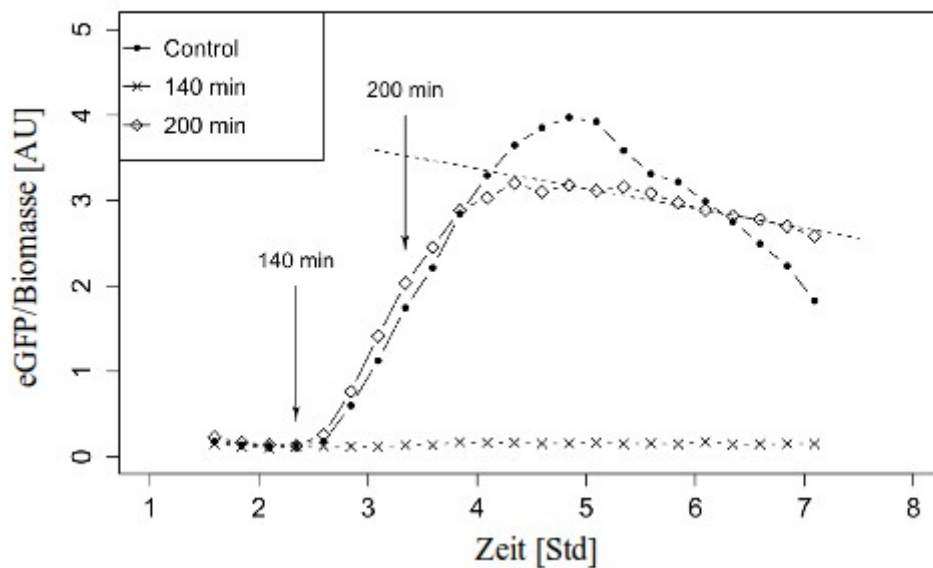


Abb. 50: Analyse der eGFP-Produktion und dessen Abbau im pSSBm85 plasmidhaltigen Stamm von *B. megaterium*. Die Induktion der eGFP-Genexpression wurde nach 140 min vorgenommen. Für diese *in vivo* Stabilitätsbestimmung von eGFP, wurde durch die Zugabe von Chloramphenicol nach 140 min und 200 min (Pfeile) die Proteinbiosynthese gehemmt. A) Nach der Zugabe von Chloramphenicol wurde das Bakterienwachstum gestoppt, was an einer früheren stationären Phase erkennbar ist. B) Als Reaktion auf Chloramphenicol wurde auch die Produktion von eGFP gestoppt. Die gefittete Kurve des Abbaus von eGFP ergab *in vivo* eine Halbwertszeit von 9,5 Stunden.

Zwischen Wachstumsrate, Plasmidverteilung und eGFP-Produktion wurde ein komplexer Zusammenhang vermutet, so dass alle Faktoren zu unterschiedlichen Zeiten nach der Induktion mit Xylose bestimmt wurden (Abb. 49). Zunächst konnte bei einer früheren Induktion weniger Biomasse, dafür mehr eGFP detektiert werden. Je später eine Kultur induziert wurde, desto weniger eGFP konnte diese bei steigender Biomasse produzieren,

Das Ausbilden zweier stabiler Subpopulationen mit unterschiedlichem Produktionsverhalten kann durch eine asymmetrische Plasmidverteilung erklärt werden, was insbesondere für die frühen Zeitpunkte nach Induktion zum Tragen kommt (Abb. 49). In diesem Fall ist die Schwankung von eGFP höher, was den Schluss einer stochastischen Plasmidverteilung zulässt. Die Produktion der Biomasse geht dabei mit weniger Ausbeute an rekombinantem Protein einher. Zusammenfassend beeinflusst eine ungleiche Verteilung der Plasmide sowohl die Populationsdynamik, als auch die Ausbeute an rekombinantem Protein.

Dieses Resultat ist nicht mit der bekannten Segregationsinstabilität durch den vollständigen Verlust des Plasmids (Summers, 1996) in Einklang zu bringen. Im Gegenteil, das Plasmid bleibt in einem “Plasmid Loss” Versuch nach Bron und Luxen, 1985 auch ohne Selektionsdruck über mehrere Stunden stabil in *B. megaterium* bestehen (siehe Kap. 3.8). Des Weiteren wird diese Art von Verlust ausschließlich unter nicht selektiven Bedingungen beobachtet. Kürzlich konnte ein Zusammenhang zwischen Genexpression und Plasmidsegregation für das ColE1-Plasmid (enthält das Gen für das Gift Colicin, das auf andere Bakterien wirkt) in *E. coli* gezeigt werden (Popov *et al.*, 2011). Es bestand hier Grund zur Annahme, dass Transkription und Translation die Plasmidreplikation beeinflusst, die daraufhin zu einer geringen Kopienzahl führt und schließlich zum Verlust des Plasmids. Ausgehend von den hier erarbeiteten Resultaten konnte allerdings kein solcher Plasmidverlust festgestellt werden, was sich mit bestehenden Arbeiten deckt (Meinhardt *et al.*, 1989).

Möglicherweise könnte unter Produktionsbedingungen die Transkriptions- und Translationsmaschinerie einen Einfluss auf die Verteilung der Plasmide in *B. megaterium* ausüben. Bei einer anderen Studie in *E. coli* konnte gezeigt werden, dass Plasmide mit erhöhter Transkription in unmittelbarer Nähe der Zellpole zu finden sind, wo eine hohe Transkriptionsaktivität herrscht (Sánchez-Romero, 2012).

Integriert man diese Ergebnisse in ein Modell (Abb. 51) und geht zunächst von einem Zustand der gleichmäßigen Plasmidverteilung aus, erzeugen alle Zellen eGFP. Vor der Zellteilung bleibt dabei ein Großteil der Plasmide an den Zellpolen, was zu Tochterzellen mit hohen und niedrigen Konzentrationen an Plasmid führt. Im Gegenzug dazu beeinflusst diese unterschiedliche Zahl der Kopien die Dosis an Genen, was sich direkt auf die Produktion rekombinanter Proteine auswirkt. Als Konsequenz entstehen Zelllinien, die aus Zellen mit hoher und niedriger Produktionsleistung bestehen. In Kombination mit den roten Zellspitzen am Ende einer wachsenden Zellkette spiegelt sich dieses Muster direkt im Modell einer wachsenden Zellkette wieder (Abb. 51 und Film 2 Anhang 8.1).

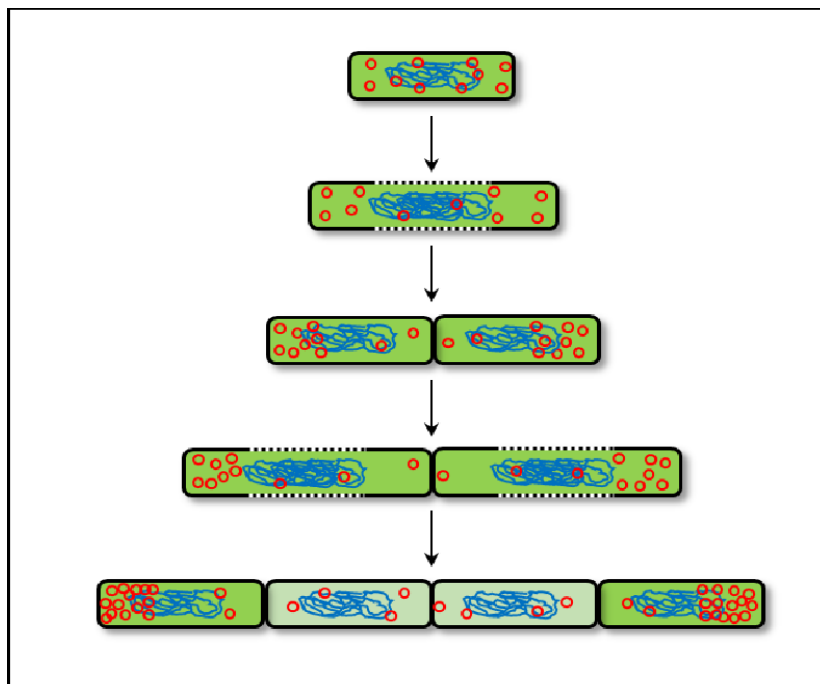


Abbildung 51: Modell zur Plasmidverteilung unter Annahme gleichmäßiger Distribution. Der Großteil der Plasmide bleibt an den Polen hängen und sorgt weiterhin für eine Produktion der rekombinanten Proteine. Aus diesem Grund entstehen sowohl Zelllinien mit hoher, als auch mit geringer Ausbeute. Innerhalb einer wachsenden Zellkette verbleiben die produktionsstärksten Zellen am Rand.

3.7 Microarray

Wachsende Mikrokolonien von *B. megaterium* zeigten mit dem Plasmid pKMMBm2, ähnlich wie der pSSBm85 Stamm, eine phänotypische Heterogenität auf Einzelzellenebene (Abb. 52). Um weitere potentielle Gene zu charakterisieren, die an diesem Geschehen direkt oder indirekt beteiligt sind, wurde ein zeitreihenabhängiges Microarrayexperiment mit dem Stamm pKMMBm2 durchgeführt und 30, 60 sowie 120 min nach Induktion das Expressionsprofil analysiert. Dazu wurden zunächst Wachstumskurven vom Wildtyp sowie den Stämmen pSSBm85 und pKMMBm2 angefertigt (Abb. 53). Diese zeigten nur einen minimalen Unterschied zwischen dem Wildtypstamm DSM 319 und seinen plasmidtragenden Stämmen. Insgesamt betrachtet wuchsen die Stämme, die ein Plasmid trugen etwas schlechter, als der plasmidfreie Wildtypstamm. Des Weiteren konnte nach der Induktion mit 0,5 % Xylose kein signifikanter Unterschied bezüglich des Wachstums der Batch Kulturen festgestellt werden.

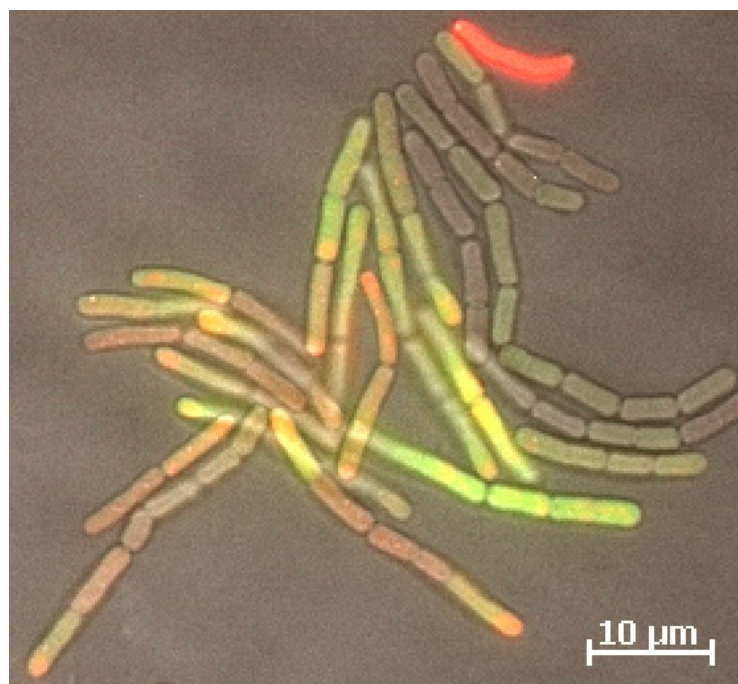


Abbildung 52: Populationsheterogenität einer wachsenden Mikrokolonie des Stammes pKMMBm2 in DSM 319. Visualisiert sind hier eGFP und XylR-mCherry.

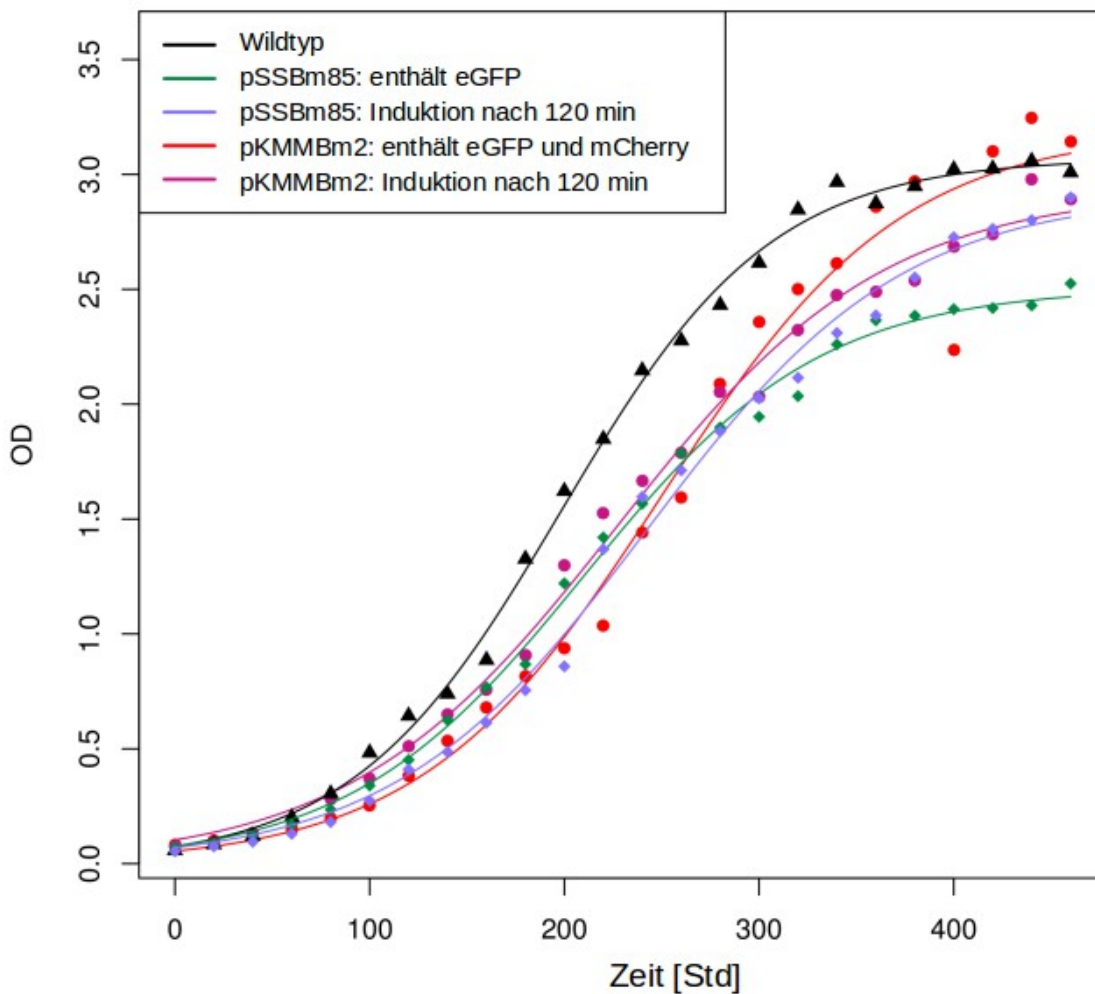


Abbildung 53: Wachstumskurven diverser Batch Kulturen von *Bacillus megaterium* mit den Plasmiden pSSBm85 und pKMMBm2 im Vergleich zum Wildtypstamm (schwarze Kurve). Die plasmidtragenden Stämme wurden nach 2 Std induziert. Als Kontrolle diente sowohl die unbehandelte Probe, als auch der Wildtyp (siehe Legende).

Die Qualität der RNA, die über den RIN-Wert bestimmt wurde, erreichte mindestens einen Wert von 8,0 (siehe Kap. 2.7.2). Die nach dem “Two-Color Microarray-Based Gene Expression Analysis Protocol“ von Agilent Technologies hybridisierte und fragmentierte RNA wurde mit dem ULS-Fluorescent Labeling Kit von Kretech mit Cy3 und Cy5 markiert.

Insgesamt konnten innerhalb der Microarrayanalyse 433 differentiell exprimierte Gene aufgezeigt werden, von denen 218 induziert und 215 reprimiert waren. Als Kriterium diente ein Expressionsverhältnis R von \pm dreifach induziert (entspricht $|\log_2(R)| > 1,59$). In Tabelle 11 wurde nach der \log_2 -Transformation die relative Expression der 50 am stärksten induzierten Gene dargestellt. Dabei unterlagen die generierten Werte diverser

Sonden für eGFP minimalen Schwankungen. Auf die verbleibenden Gene wird später genauer eingegangen. Die Genliste und das dazugehörige R-Skript zur Datenauswertung sind im Anhang aufgeführt.

Eine Übersicht der stärksten 50 differentiell exprimierten Gene kann aus dem Clusterdendrogramm in Abbildung 54 gewonnen werden. Diese Heatmap verdeutlicht den Mittelwert, der am stärksten vorkommenden Genexpressionsunterschiede nach Induktion.

Tabelle 11: Übersicht der 50 am stärksten induzierten und reprimierten Gene während der Induktion der Proteinproduktion. Als Kriterium diente die Summe des Betrags der logarithmierten Expressionsverhältnisse (logFC) nach 30, 60 und 120 Minuten ($\sum |\log_2(FC)|$). Dargestellt wurden die systematische Gennummer (ORF ID), Genname mit Beschreibung und die logFC Werte mit den ermittelten P-Werten (Irrtumswahrscheinlichkeiten) zu den gemessenen Zeitpunkten.

ORF ID	Name	Beschreibung	logFC30	PVal30	logFC60	PVal60	logFC120	PVal120
BMD_0684	argF	ornithine carbamoyltransferase	3,905	2,44E-01	3,351	2,15E-04	2,573	3,90E-05
BMD_4657	maf	septum formation protein Maf	2,335	2,44E-01	3,746	2,84E-05	2,932	8,21E-05
BMD_2932		hypothetical protein	1,676	2,44E-01	3,360	7,30E-05	3,639	3,36E-05
BMD_1858	xylA	xylose isomerase	2,379	2,44E-01	3,997	6,88E-06	2,063	4,54E-05
BMD_2605	cblX	sirohydrochlorin cobaltochelata	-2,358	2,44E-01	-3,750	1,18E-04	-2,574	5,82E-05
BMD_4656	radC	DNA repair protein RadC	2,344	2,44E-01	3,748	1,33E-05	2,329	3,36E-05
BMD_1450	cspD	cold shock protein	1,542	2,44E-01	3,184	4,46E-05	3,522	6,63E-05
BMD_4063	gabT	4-aminobutyrate transaminase	-2,338	2,44E-01	-3,536	1,26E-04	-2,712	3,47E-04
BMD_2208		general stress protein 17M	2,126	2,44E-01	3,344	7,30E-05	3,109	4,90E-05
GFP_62	GFP_62	Unknown	1,161	2,72E-01	2,732	7,27E-05	3,822	9,04E-05
BMD_1857	xylR	xylose repressor	1,767	2,44E-01	3,034	2,84E-05	3,511	5,51E-05
BMD_2897		conserved hypothetical protein	1,627	2,44E-01	2,268	4,27E-04	3,842	4,50E-05
GFP_171	GFP_171	Unknown	1,025	3,32E-01	2,092	2,70E-04	3,849	8,51E-05
BMD_1557		glucose starvation-inducible protein B	1,325	2,44E-01	2,242	1,10E-04	3,673	3,36E-05
BMD_4776	argG	argininosuccinate synthase	2,864	2,44E-01	3,213	2,24E-04	1,569	7,52E-05
BMD_2603	cblC	precorrin-8X methylmutase CblC	-2,144	2,44E-01	-3,349	1,66E-04	-2,241	7,50E-05
BMD_4713	lysC	aspartate kinase	0,386	2,46E-01	3,129	7,29E-03	3,078	3,56E-05
BMD_3013		hypothetical protein	0,663	2,50E-01	1,860	5,18E-04	3,726	6,56E-05
BMD_4775	argH	argininosuccinate lyase	2,992	2,44E-01	2,965	3,96E-04	1,396	9,85E-05
GFP_203	GFP_203	Unknown	0,902	3,16E-01	2,227	1,57E-04	3,565	1,03E-04
BMD_3145		hypothetical protein	1,963	2,44E-01	3,203	4,46E-05	2,172	1,74E-04
BMD_0683	carB	carbamoyl-phosphate synthase, large subunit	2,973	2,44E-01	2,782	3,39E-04	1,405	4,97E-05
BMD_0904		hypothetical protein	1,170	2,45E-01	2,490	1,05E-03	3,240	4,54E-05
BMD_0681	argD	acetylornithine aminotransferase	2,949	2,44E-01	2,561	4,48E-04	1,531	7,18E-05
BMD_2606	cblH	precorrin 3 methylase	-2,010	2,44E-01	-3,041	2,25E-04	-2,051	1,91E-04
BMD_1682	cspA	cold shock protein	1,510	2,44E-01	2,542	9,82E-05	2,852	3,36E-05
GFP_229	GFP_229	Unknown	0,753	3,33E-01	1,881	3,58E-04	3,375	1,97E-04
BMD_2600	cblL	precorrin-2 C20-methyltransferase	-1,915	2,44E-01	-3,074	1,37E-04	-1,799	5,16E-05
BMD_0680	argB	acetylglutamate kinase	2,884	2,44E-01	2,606	5,38E-04	1,087	3,53E-04
BMD_0148	rpsH	30S ribosomal protein S8	-0,973	2,50E-01	-2,000	6,00E-04	-3,189	3,90E-04
BMD_3693	bioF	8-amino-7-oxononanoate synthase	1,671	2,44E-01	3,000	2,58E-03	1,879	8,05E-03
BMD_2599	cblF	precorrin-4 C11-methyltransferase	-1,919	2,44E-01	-2,908	1,03E-04	-1,818	8,69E-05
BMD_2910		conserved hypothetical protein	-1,170	2,44E-01	-2,729	2,84E-05	-2,536	5,16E-05
BMD_0958	treC	alpha, alpha-phosphotrehalase	-1,880	2,44E-01	-2,817	1,89E-04	-1,972	6,93E-04
BMD_2601	cblET	CblET protein	-1,889	2,44E-01	-2,923	1,17E-04	-1,668	2,09E-04

Populationsheterogenität während der heterologen Proteinproduktion in *Bacillus megaterium*

GFP_145	GFP_145	Unknown	0,670	3,60E-01	1,393	6,00E-04	3,350	2,09E-04
BMD_0135	rplD	50S ribosomal protein L4	-0,692	2,60E-01	-1,576	6,78E-04	-3,273	1,96E-04
BMD_0682	carA	carbamoyl-phosphate synthase, small subunit	2,790	2,44E-01	2,012	3,49E-04	1,720	4,93E-05
BMD_2285	nprM	Bacillolysins precursor (neutral protease)	1,180	2,44E-01	2,705	2,84E-05	2,432	8,60E-05
BMD_4636	nadA	quinolinate synthetase complex, A subunit	0,235	3,71E-01	2,985	5,55E-03	2,297	2,80E-04
BMD_2602	cbiD	cobalamin biosynthesis protein CbiD	-1,823	2,44E-01	-2,863	2,49E-04	-1,579	4,27E-04
BMD_3867	nucB	sporulation-specific extracellular nuclease	1,748	2,44E-01	2,584	5,30E-05	2,151	5,10E-05
BMD_0957	scrA	PTS system trehalose-specific EIIBC component	-1,815	2,44E-01	-2,686	1,07E-04	-1,828	4,09E-04
BMD_1178		conserved hypothetical protein	1,733	2,44E-01	2,674	1,84E-04	1,878	7,39E-04
BMD_3090		conserved hypothetical protein	1,625	2,44E-01	2,341	1,14E-04	2,363	1,51E-03
BMD_4638	nadB	L-aspartate oxidase	0,483	2,54E-01	2,593	1,15E-03	2,522	3,36E-05
BMD_4416	artM	arginine ABC transporter, ATP-binding protein ArtM	2,190	2,44E-01	2,498	5,76E-04	1,523	4,15E-04
SPYoch	SPYoch	Unknown	0,430	3,61E-01	0,669	2,48E-01	3,342	4,54E-05
BMD_5137	atpH	ATP synthase F1, delta subunit	-1,412	2,44E-01	-2,221	3,54E-05	-2,524	7,55E-05
BMD_4049		Transporter (Major facilitator Superfamily)	-1,650	2,44E-01	-2,517	3,47E-04	-2,053	1,66E-04

Innerhalb der hochregulierten Gruppe waren insbesondere Operons, wie die der Argininbiosynthese, *gfp*, *xylR* und *xylA* involviert, während in der runterregulierten Gruppe solche der zentralen Stoffwechselwege wie Cobalamin- und Precorrinbiosynthese zu finden waren. Arginin ist u. a. beim Aminosäureabbau involviert und wird während Glucosemangel in Zellen von *B. licheniformis* hochreguliert (Voigt et al., 2014). Der Argininabbau steht dabei ebenso unter der Kontrolle der Katabolitrepression, die durch Glucose reprimiert wird (Laishley and Bernlohr, 1968). Bei *B. subtilis* steht der Argininabbau ebenfalls unter der Kontrolle von CcpA (Belitsky et al., 2004). Über den Aminosäureabbau wird ein Großteil des Oxalacetats geliefert, das für die Gluconeogenese benötigt wird. Die erhöhte Expression der Gene des Arginin-Operons deutet daher höchstwahrscheinlich auf Hungerbedingungen während der heterologen Proteinproduktion hin.

Grundsätzlich konnten 30 min nach Induktion 45 regulierte Gene ermittelt werden, 1 Std nach Xylosezugabe 269 Gene und 2 Std danach wurden 298 Gene aufgedeckt (Abb. 55A). Somit war ein deutlicher Anstieg an direkt und indirekt beteiligten Genen nach Induktion der Proteinproduktion zu verzeichnen. Darüber hinaus verdeutlicht das Venn-Diagramm in Abbildung 55B die Schnittmenge, der jeweils differentiell exprimierten Gene zu den unterschiedlichen Zeitpunkten. Auffällig war, dass alle zu Beginn regulierten Gene ebenfalls nach 1 Std detektiert werden konnten, wohingegen nur 28 bei der darauf folgenden Probenahme gefunden wurden. Ebenso gab es 118 Gene, die

ausschließlich zum ersten Zeitpunkt nach Induktion differentiell exprimiert wurden und 164 Gene zum zweiten Zeitpunkt. Die Schnittmenge der letzten beiden Zeitpunkte lag bei 134 Genen. Die Gene ließen sich demzufolge sowohl in eine unmittelbare und lang anhaltende Reaktion, als auch in eine spät temporäre Reaktion gruppieren.

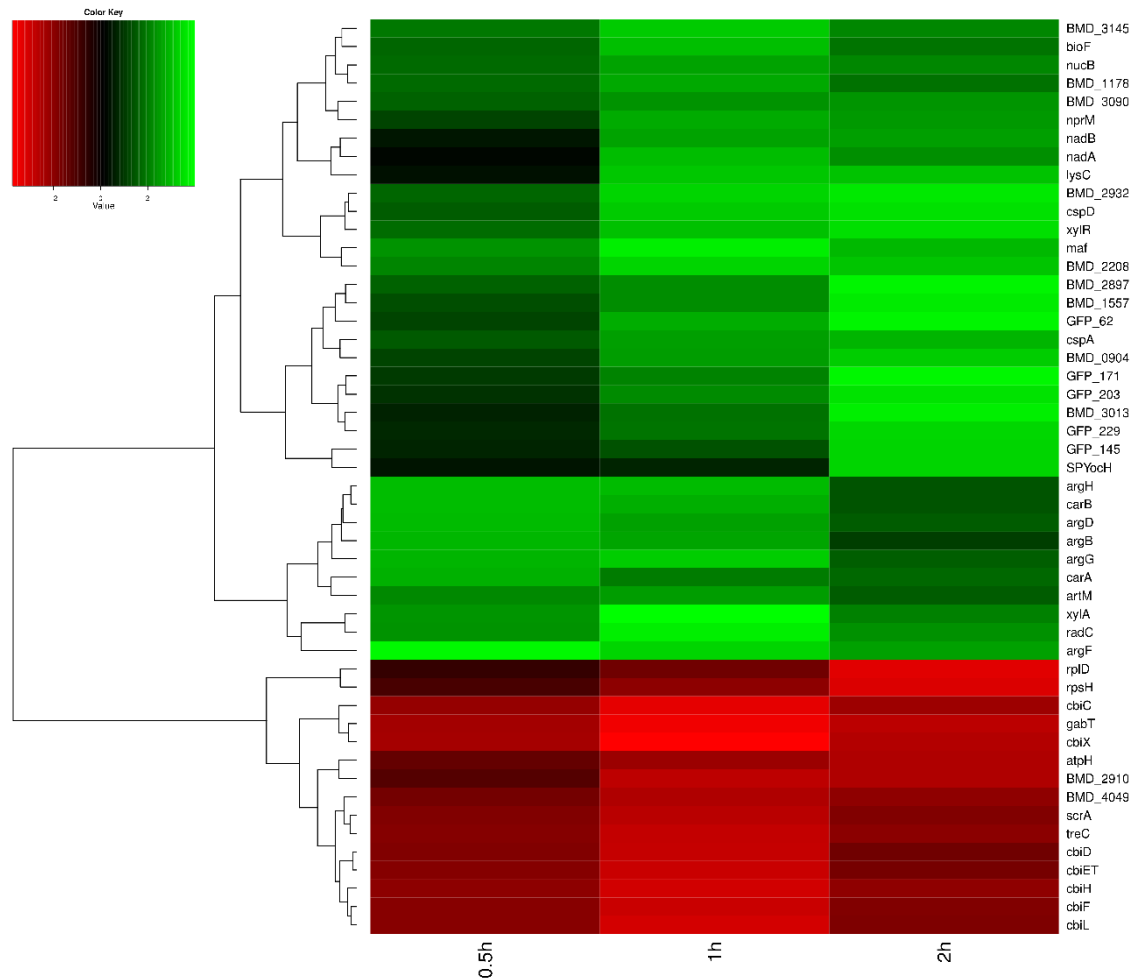


Abbildung 54: Clusterdendrogramm des Microarrays mit Farbschlüssel. In dieser Abbildung wurden die ersten 50 Gene mit den höchsten Expressionsunterschieden in ihrem zeitlichen Expressionsmuster geclustert. Die grüne Farbe ist mit einer induzierten Expression verknüpft, wohingegen die Farbe Rot mit einer Repression assoziiert ist. Als Kriterium wurde hier die Summe des Betrags der Expressionsverhältnisse (FC) der drei Zeitpunkte t verwendet ($\sum |\log_2(FC)|$).

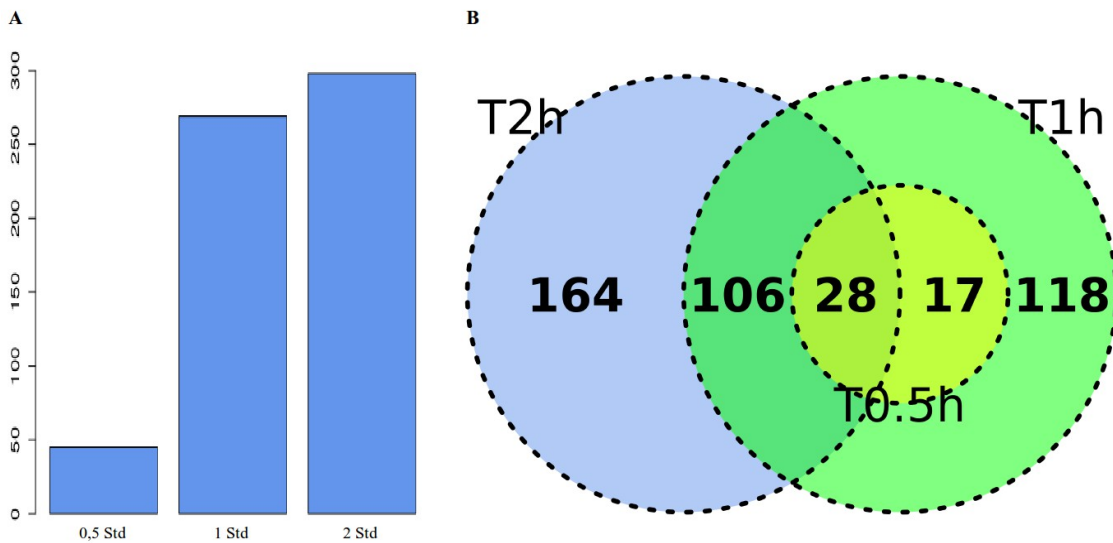


Abbildung 55: Statistik der differentiell exprimierten Gene. A) Das Balkendiagramm beschreibt die Anzahl der differentiell exprimierten Gene zu den Zeitpunkten 30 min, 60 min sowie 120 min nach Induktion. B) Das Venn-Diagramm verdeutlicht die Schnittmenge der jeweiligen Gene. Hellgrün zeigt die Gene zum Zeitpunkt der ersten Probenahme, die Farbe Dunkelgrün zur zweiten Entnahme sowie Blau zum letzten Zeitpunkt.

Auch hier sollte eine Tabelle Aufschluß über die 45 Gene geben, die zu Beginn induziert wurden und 120 min nach Induktion immer noch exprimiert waren da man hier die Kandidaten vermuten würde, die in direkter Art und Weise vom Produktionsprozess beeinflusst wurden (28 Gene im blauen Überlappungsbereich des Venn-Diagramms in Abb. 55B entsprechen dem ersten Teil von Tabelle 12). Bei den restlichen 17 Genen (Tab. 12) handelte es sich ausschließlich um jene, die zu einem frühen Zeitpunkt nach der Induktion (30 min und 60 min) differentiell exprimiert waren (siehe auch Venn-Diagramm Abb. 55B),

Auffällig ist die Tatsache, daß eGFP erst zu einem späteren Zeitpunkt induziert wird, da es unmittelbar 30 min nach Xylosezugabe fehlt. Auch waren nur zu einem frühen Zeitpunkt die Gene des Arg-Operons von Bedeutung, wohingegen zu einem späteren Zeitpunkt lediglich *argF* noch detektiert werden konnte. Ist genügend Arginin vorhanden, ist der Gehalt an *argF* gering, da es sich hierbei um eine Endproduktrepression handelt. Wird jedoch gleichzeitig vermehrt Glutamin über die Induktion von *carA* (Carbonyl-Phosphate Synthase) abgebaut, entsteht durch ArgF wieder Arginin.

Darüber hinaus wird das Operon, das an der Biosynthese von Cobalamin beteiligt ist,

exprimiert. Zum einen ist Vitamin B12 wichtig für das Zellwachstum und zum anderen bei der Übertragung von Methylgruppen. Dabei entsteht über die Methionin-Synthase Methionin. Zudem kontrolliert das Gen *maf* unter anderem die Zellteilung (Butler *et al.*, 1993). Wie bereits erwähnt, wurde *ftsA* nach andauernder Induktion immer mehr reprimiert und hemmte das Zellwachstum, was die Resultate der langgestreckten Zellen in den Filmen belegten.

Bei Sigma 54 handelt es sich um einen Sigma-Faktor, der bei Stickstoffmangel gebildet wird. Er bewirkt die Expression der Glutaminsynthetase, die das Schlüsselenzym für die Stickstoff-Assimilierung darstellt.

Tabelle 12: Übersicht der überlappend induzierten und reprimierten Gene nach allen Zeitpunkten (oberer Bereich) sowie zu den Zeitpunkten 30 und 60 Minuten (unterer Bereich) gemäß dem Venn-Diagramm in Abbildung 54. Als Kriterium für eine Expression diente ein mindestens 3-fach induziertes oder reprimiertes Verhältnis bzw. $|\log_2(FC)| > 1,58$.

ORF ID	Name	Beschreibung	logFC30	PVal30	logFC60	PVal60	logFC120	PVal120
BMD_0684	argF	ornithine carbamoyltransferase	3,905	2,44E-01	3,351	2,15E-04	2,573	3,90E-05
BMD_4657	maf	septum formation protein Maf	2,335	2,44E-01	3,746	2,84E-05	2,932	8,21E-05
BMD_2932		hypothetical protein	1,676	2,44E-01	3,360	7,30E-05	3,639	3,36E-05
BMD_1858	xylA	xylose isomerase	2,379	2,44E-01	3,997	6,88E-06	2,063	4,54E-05
BMD_2605	cblX	sirohydrochlorin cobaltochelataase	-2,358	2,44E-01	-3,750	1,18E-04	-2,574	5,82E-05
BMD_4656	radC	DNA repair protein RadC	2,344	2,44E-01	3,748	1,33E-05	2,329	3,36E-05
BMD_4063	gabT	4-aminobutyrate transaminase	-2,338	2,44E-01	-3,536	1,26E-04	-2,712	3,47E-04
BMD_2208		general stress protein 17M	2,126	2,44E-01	3,344	7,30E-05	3,109	4,90E-05
BMD_1857	xylR	xylose repressor	1,767	2,44E-01	3,034	2,84E-05	3,511	5,51E-05
BMD_2897		conserved hypothetical protein	1,627	2,44E-01	2,268	4,27E-04	3,842	4,50E-05
BMD_2603	cblC	precorrin-8X methylmutase CblC	-2,144	2,44E-01	-3,349	1,66E-04	-2,241	7,50E-05
BMD_3145		hypothetical protein	1,963	2,44E-01	3,203	4,46E-05	2,172	1,74E-04
BMD_2606	cblH	precorrin 3 methylase	-2,010	2,44E-01	-3,041	2,25E-04	-2,051	1,91E-04
BMD_2600	cblL	precorrin-2 C20-methyltransferase	-1,915	2,44E-01	-3,074	1,37E-04	-1,799	5,16E-05
BMD_3693	bioF	8-amino-7-oxononanoate synthase	1,671	2,44E-01	3,000	2,58E-03	1,879	8,05E-03
BMD_2599	cblF	precorrin-4 C11-methyltransferase	-1,919	2,44E-01	-2,908	1,03E-04	-1,818	8,69E-05
BMD_0958	treC	alpha, alpha-phosphotrehalase	-1,880	2,44E-01	-2,817	1,89E-04	-1,972	6,93E-04
BMD_2601	cblET	CblET protein	-1,889	2,44E-01	-2,923	1,17E-04	-1,668	2,09E-04
BMD_0682	carA	carbamoyl-phosphate synthase, small subunit	2,790	2,44E-01	2,012	3,49E-04	1,720	4,93E-05
BMD_3867	nucB	sporulation-specific extracellular nuclease	1,748	2,44E-01	2,584	5,30E-05	2,151	5,10E-05
BMD_0957	scrA	PTS system trehalose-specific EIIBC component	-1,815	2,44E-01	-2,686	1,07E-04	-1,828	4,09E-04
BMD_1178		conserved hypothetical protein	1,733	2,44E-01	2,674	1,84E-04	1,878	7,39E-04
BMD_3090		conserved hypothetical protein	1,625	2,44E-01	2,341	1,14E-04	2,363	1,51E-03
BMD_4049		Transporter (Major facilitator Superfamily)	-1,650	2,44E-01	-2,517	3,47E-04	-2,053	1,66E-04
BMD_3217	yclP	putative ferrichrome ABC transporter, ATP-binding protein	-1,653	2,44E-01	-2,460	1,18E-04	-2,110	6,52E-05
BMD_4062		Sigma-54 interaction domain	-1,762	2,44E-01	-2,441	4,86E-04	-1,637	2,14E-03
BMD_4061		succinate-semialdehyde dehydrogenase [NADP+]	-1,629	2,44E-01	-2,309	3,27E-04	-1,944	7,71E-04
BMD_0894		conserved hypothetical protein	1,767	2,44E-01	2,087	5,10E-04	1,831	1,34E-04
BMD_4776	argG	argininosuccinate synthase	2,864	2,44E-01	3,213	2,24E-04	1,569	7,52E-05
BMD_4775	argH	argininosuccinate lyase	2,992	2,44E-01	2,965	3,96E-04	1,396	9,85E-05

Populationsheterogenität während der heterologen Proteinproduktion in *Bacillus megaterium*

BMD_0683	carB	carbamoyl-phosphate synthase, large subunit	2,973	2,44E-01	2,782	3,39E-04	1,405	4,97E-05
BMD_0681	argD	acetylornithine aminotransferase	2,949	2,44E-01	2,561	4,48E-04	1,531	7,18E-05
BMD_0680	argB	acetylglutamate kinase	2,884	2,44E-01	2,606	5,38E-04	1,087	3,53E-04
BMD_2602	cbiD	cobalamin biosynthesis protein CbiD	-1,823	2,44E-01	-2,863	2,49E-04	-1,579	4,27E-04
BMD_4416	artM	arginine ABC transporter, ATP-binding protein ArtM	2,190	2,44E-01	2,498	5,76E-04	1,523	4,15E-04
BMD_1479		conserved hypothetical protein	1,705	2,44E-01	2,896	3,64E-05	1,182	5,76E-04
BMD_2604	cbiJ	precorrin-6x reductase	-1,745	2,44E-01	-2,807	2,15E-04	-1,366	5,22E-04
BMD_1859	xylB	xylulokinase	1,689	2,44E-01	2,557	5,70E-05	0,958	6,63E-04
BMD_0679	argJ	arginine biosynthesis bifunctional protein ArgJ	2,521	2,44E-01	1,683	1,92E-03	0,587	1,67E-03
BMD_5000	yfiY	putative ferrichrome import ABC transporter, ferrichrome-binding protein	-1,611	2,44E-01	-2,053	3,69E-04	-1,547	1,82E-03
BMD_0945		putative 4-aminobutyrate transaminase	2,138	2,44E-01	1,819	4,86E-04	1,037	1,53E-04
BMD_0678	argC	N-acetyl-gamma-glutamyl-phosphate reductase	2,284	2,44E-01	1,602	1,20E-03	0,589	2,48E-03
BMD_4032		FAD-dependent dehydrogenase	1,604	2,44E-01	2,064	7,27E-05	0,732	5,44E-03
BMD_4417	artQ	arginine ABC transporter, permease protein ArtQ	1,806	2,44E-01	1,922	1,14E-03	0,640	3,35E-03
BMD_4418	artP	arginine ABC transporter, arginine-binding protein ArtP	1,678	2,45E-01	1,623	2,12E-03	0,566	9,50E-03

Prinzipiell wurden einige wichtige, zentrale Stoffwechselwege blockiert. Dies soll anhand des Citratzyklus (TCA) in Abbildung 56 verdeutlicht werden. Die meisten der dafür benötigten Gene wurden reprimiert (rot), während nur wenige Gene exprimiert (blau) wurden. Die Pyruvatcarboxylase würde die Addition von Kohlenstoffdioxid an Pyruvat katalysieren, woraufhin Oxalacetat entsteht, das zusammen mit Acetyl-CoA den ersten Schritt des TCA einleitet und als Resultat Citrat entsteht. Ein weiteres Beispiel stellt Abbildung 57 dar. Hier wurde das Enzym Acetyl-CoA-Carboxylase reprimiert, das den ersten Schritt der Lipogenese einleitet. Offenbar herrschte dadurch ein Mangel an Fettsäuren, so dass diese nicht als Energiereserven dienen konnten. Daraus wird deutlich, dass viele zentrale Gene, die wichtig für den Energiestoffwechsel sind aktiv reprimiert wurden. Darüber hinaus nahm die Zellteilungsrate ab, da *ftsA* nach andauern- der Induktion mehr und mehr reprimiert wurde. Dieses Resultat legt die Vermutung nahe, dass sich die Zellen zwar unmittelbar äußerst produktiv zeigen, längerfristig evtl. aber verhungern. Dies wurde unterstützend durch das oben beschriebene Live/Dead-Staining (Lebend/Tot-Färbung) im halbdefinierten A5-Medium getestet und belegt.

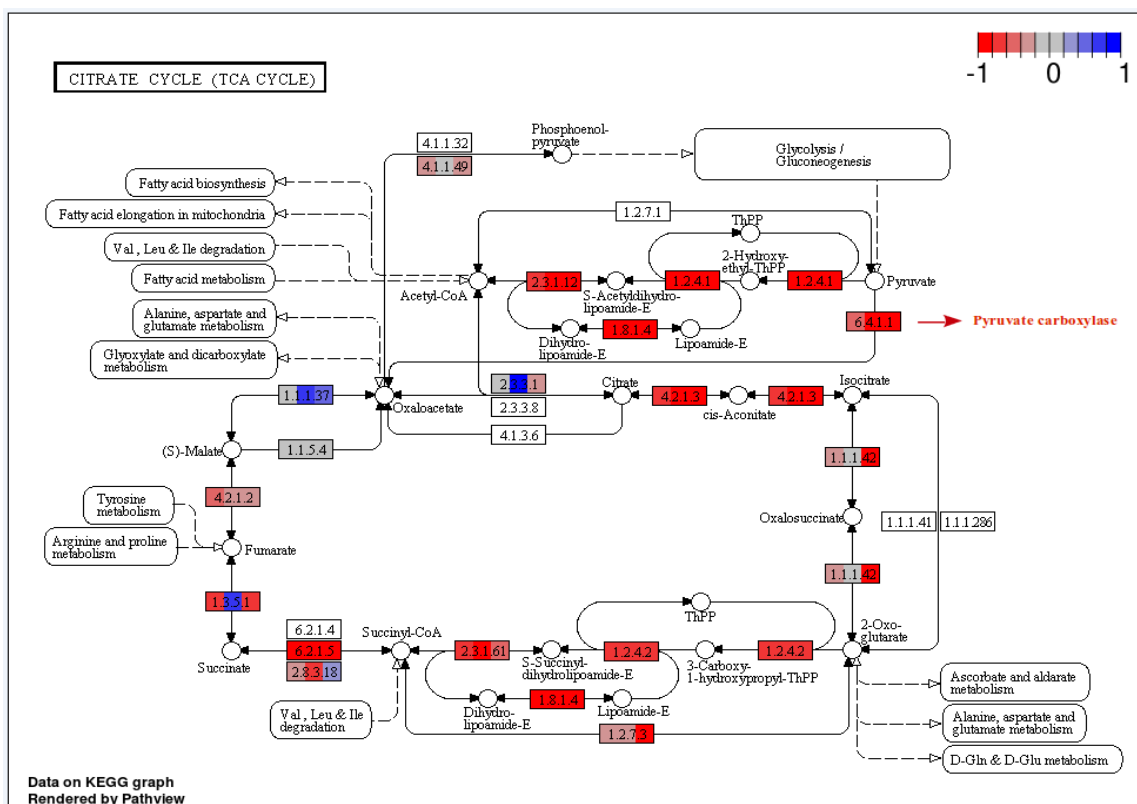


Abbildung 56: Visualisierung metabolischer Stoffwechselwege der differentiell regulierten Gene des Microarray mit Hilfe von KEGG basierten Stoffwechselkarten. Die meisten Gene des Citratzyklus (TCA) wurden reprimiert (rot dargestellt). Diese Beobachtung gab es auch bei anderen zentralen Stoffwechselwegen.

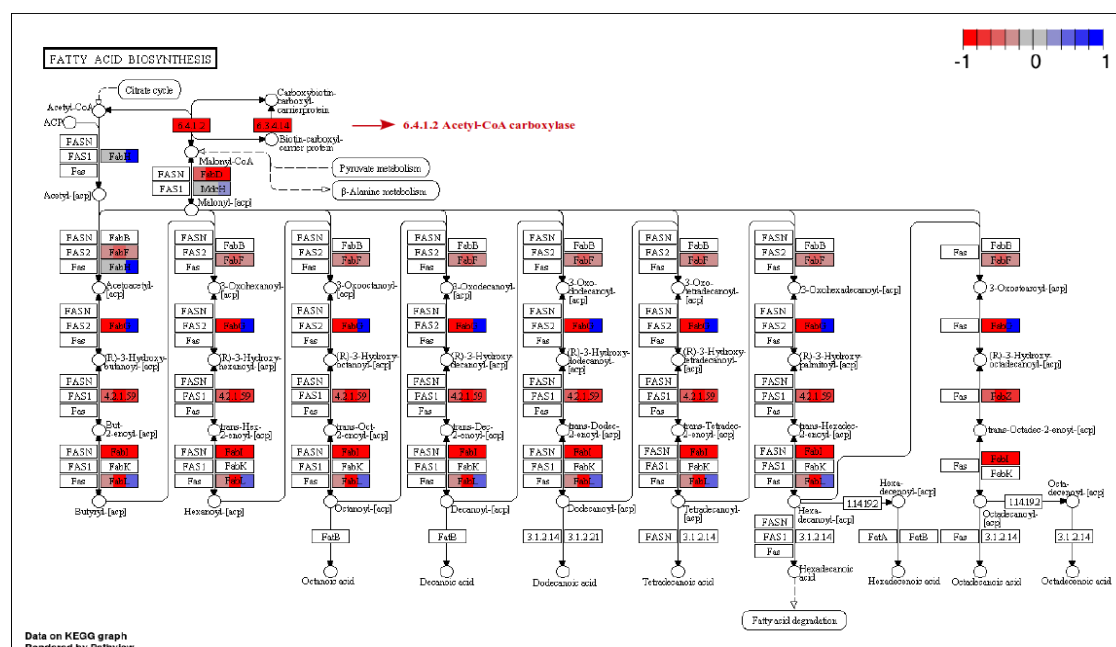


Abbildung 57: Resultat der Fettsäuresynthese aus dem Microarray basierend auf KEGG metabolische Stoffwechselkarten. Die meisten Gene waren zunächst reprimiert (rot dargestellt), wohingegen 2 Std nach Induktion eine Expression der reduktiven Elongation (blauer Bereich von FabG – Beta-Ketoacyl-Acyl Carrier Protein Reductase) stattfand.

3.8 Wachstumskurven und mikroskopische Populationsheterogenität im Vergleich

Die Analyse der verwendeten Stämme bezüglich des Wachstums sowie fluoreszenzmikroskopische Aufnahmen sollten Gemeinsamkeiten und Unterschiede aufzeigen, die als Grundlage diverser Untersuchungen diene.

3.8.1 Wachstum des Wildtyps im Vergleich zu den plasmidtragenden Stämmen

Die mit dem BioLector gemessene Wachstumskurven in Abbildung 58 zeigten ohne Induktion (Abb. 58A) nur einen minimalen Unterschied zwischen dem Wildtypstamm DSM 319 und seinen plasmidtragenden Stämmen. Insgesamt betrachtet wuchsen diese, die ein Plasmid trugen etwas schlechter, als der plasmidfreie Wildtypstamm. Ein deutlicher Unterschied konnte bei pKMMBm5 festgestellt werden, da dieser Stamm bereits von Anfang an – ohne zusätzliche Xylosezugabe – eGFP produzierte. Induzierte man die Kulturen in der exponentiellen Wachstumsphase nach 150 min mit Xylose, so wuchs der Wildtypstamm erneut besser, als die plasmidtragenden Derivate (Abb. 58B). Allerdings nahm das Wachstum in allen Fällen signifikant ab, lediglich die *xylR*-Mutante pKMMBm5 zeigte weiterhin eine stetige OD-Zunahme. Dieser Stamm ist bereits zu Beginn einer Kultur mit dem Stressfaktor der eGFP-Produktion vertraut. Darüber hinaus belegen Hinweise den Verdacht, dass Xylose verstoffwechselt wird und notwendige Energiereserven dieser Mutante auffüllt. Eine Umstellung mit zusätzlicher Produktionsleistung geht in diesem Fall nicht mit der Induktion einher.

Demnach belegen die Wachstumskurven ein Stagnieren unmittelbar nach der Induktion. Aus den Filmen der Fluoreszenzmikroskopie (Kap. 3.4) ging hervor, dass grüne Zellen deutlich langsamer wuchsen und sich langsamer teilten, was auch die Arraydaten (Kapitel 3.7) bestätigten, da *ftsA* nach Induktion beständig abnahm. Die Top-Produzenten verbrauchen demnach offenbar weniger Energie für Wachstum und setzen diese in Produktionsleistung um, was allerdings das Fortbestehen nicht garantiert. Diese Zellen sind durch den knappen Energiehaushalt, der in Kapitel 3.7 durch die Microarrayanalyse beschrieben wurde, nicht in der Lage, Zellen mit hoher Plasmidzahl zu generieren, was die Überlebensstrategie mittels freier Diffusion nicht garantieren und die Nachkom-

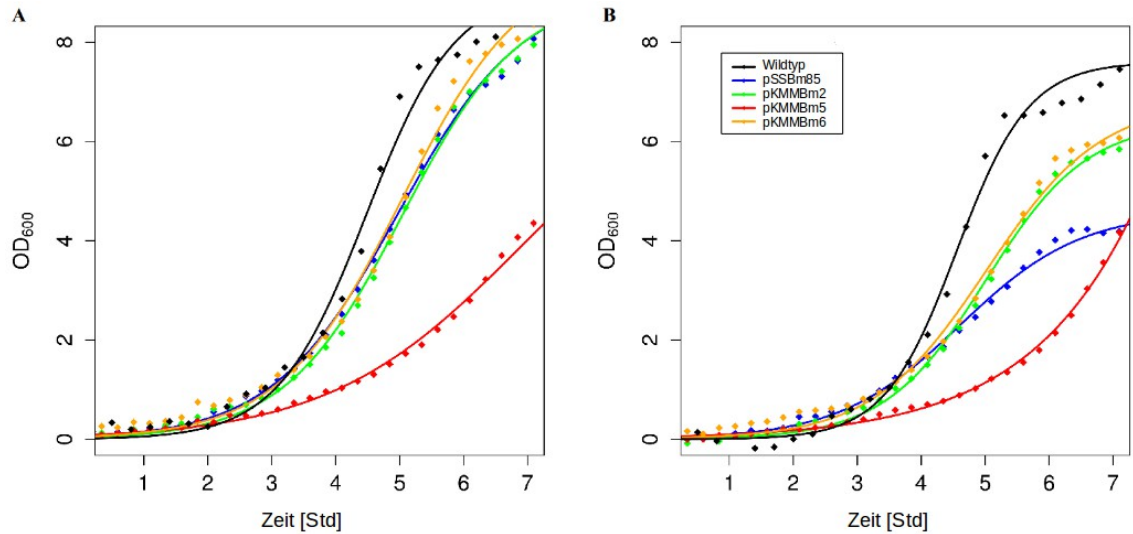


Abbildung 58: Wachstumskurven der verwendeten Stämme mit und ohne Zugabe von Xylose gemessen mit Hilfe des BioLectors. A) Ohne Zugabe von Xylose zeigt das Wachstum nur wenig Unterschied, wobei pKMMBm5 durch die sofortige Produktionsleistung von eGFP zeitlich verzögert wächst. B) Nach der Induktion mit Xylose kann ein Abfall der Populationsdichte für die Plasmide pSSBm85, pKMMBm2 und pKMMBm6 festgestellt werden, wohingegen die *xyIR*-Mutante pKMMBm5 wenig Unterschied zeigte (rote Kurve).

men sichern würde, so dass an dieser Stelle evtl. auf einen kontrollierten Mechanismus bei der Plasmidverteilung auf die Tochterzellen zurückgegriffen wird. Belegt wurde dieses Ereignis durch die FISH-Experimente (Kapitel 3.6). Im nicht induzierten Zustand bildeten sich keine Plasmidfoci, so dass eine Gleichverteilung möglich war. Eine induzierte Kultur hingegen wies deutliche Plasmidkluster auf, die eine Gleichverteilung unmöglich machten. Der XylR-Gehalt in grünen Zellen stieg durch die Transkription des Plasmids an, was die Daten der qRT-PCR in Kapitel 3.3.4 dokumentierten. Diese könnten nicht mehr abgefangen werden, da die Energie überwiegend für die Produktion von eGFP und XylR zur Verfügung gestellt wird und nicht für das Wachstum der Zellen. Zudem wurden die Kulturen permanent in der frühen exponentiellen Phase überimpft. Eine Phase, bei der gemäß den Resultaten des Durchflusszytometers in Kapitel 3.6 der Anteil der weißen Zellen nach Induktion zunächst gewaltig sank. Vermutlich kommen hier nun Hungers- und Alterserscheinungen zum Tragen, die die Rückkehr der erschöpften Kultur zur weißen Population im weiteren Verlauf bewirkten. Nachdem bei dem Versuch zur Stabilität des Plasmids in Kapitel 3.1 (Abb. 13) jedoch stets innerhalb der späten exponentiellen Phase die Kultur in neues Medium transferiert und die anderen Wachstumsphasen komplett außer Acht gelassen wurden, könnte es sich bei der

Heterogenität um eine kurze Momentaufnahme einer komplexen Situation handeln. Erstaunlich bleibt jedoch die Tatsache, wie entscheidend und damit einflussreich die Vorgeschichte bezüglich der Induktion für den weiteren Weg über lange Zeit beständig blieb.

3.8.2 Mikroskopische Fluoreszenzaufnahmen diverser plasmidtragender Stämme

Bei allen Derivaten des Wildtypstamms konnte nach der Induktion mit Xylose eine Populationsheterogenität erkannt werden (Abb. 59). Festgestellt wurde das beschriebene Phänomen einst beim eGFP enthaltenden Stamm pSSBm85 (Abb. 59A). Durch das Fusionieren von XylR mit mCherry konnte weiterhin eine Heterogenität beobachtet werden, zuzüglich der Entdeckung von Zellen, die XylR ansammelten (siehe tiefrote Zelle in Abb. 59B). Wurde XylR aus dem Plasmid entfernt, konnte keine wesentliche Änderung bezüglich der heterogenen Eigenschaft festgestellt werden (Abb. 59C). Die größte Heterogenität gemäß der FACS-Analysen in Kapitel 3.6.1 (Abb. 40) erzielte man mit pKMMBm6 (Abb. 59D und E). Dabei handelte es sich um jenen Stamm, der zehn repetitive Bindestellen für XylR enthielt. In diesem Fall war der Anteil an rein weißen Zellen nach der Induktion relativ hoch. Offensichtlich wirkte sich diese Wegtitration des Repressors zwar auf die Ausbildung von mehr weißen Zellen aus, nicht aber auf die eigentliche Heterogenität. Diese Daten widerlegen damit die anfängliche Hypothese eines Verteilungsphänomens des Repressors XylR und belegen die gezeigte ungleichmäßige Plasmidverteilung nach der Induktion.

Zum Schluß soll eine DAPI-Färbung von pSSBm85 die verschiedenen Zustände der DNA – kompakt oder spiralisiert – nach dem Induzieren veranschaulichen (Abb. 60).

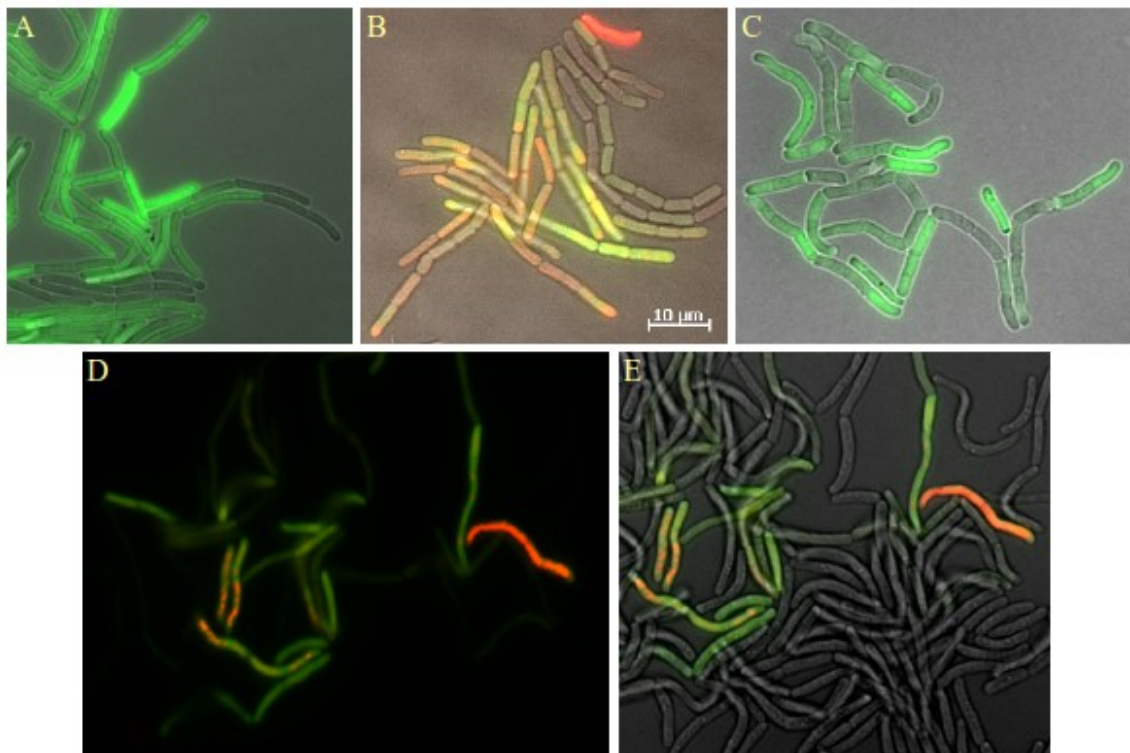


Abbildung 59: Phänotyp diverser verwendeter Stämme. Fluoreszenzmikroskopische Aufnahmen einer Kultur von *Bacillus megaterium* mit den Stämmen A) pSSBm85, B) pKMMBm2, C) pKMMBm5 sowie D) und E) pKMMBm6. Alle Stämme besitzen eine ausgeprägte Kulturheterogenität.

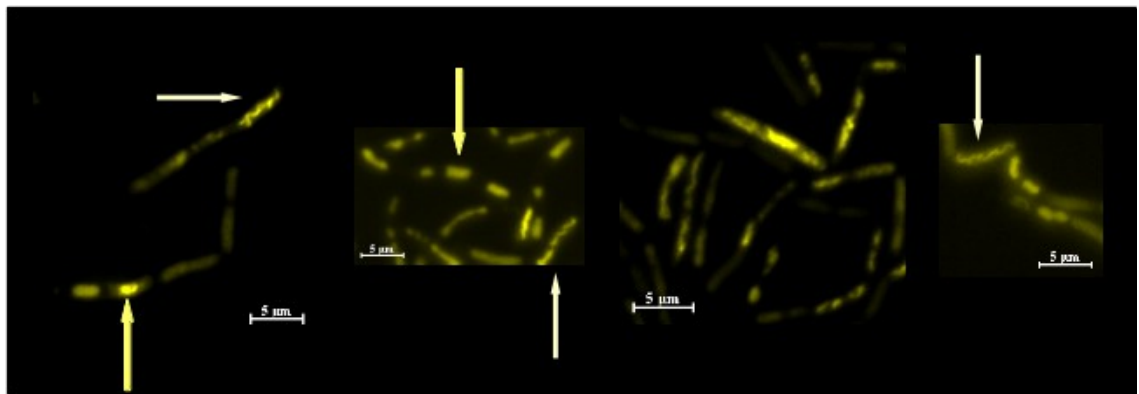


Abbildung 60: DAPI-Färbung des Stamms pSSBm85. Fluoreszenzmikroskopische Aufnahmen zeigen die DNA nach der Zugabe von Xylose als Knäuel (blassgelbe Pfeile) oder in spiralisierter Form (weiße Pfeile).

4. Zusammenfassung

4.1 Bistabilität wird offensichtlich nicht durch die positive Rückkopplung über die Xylose-Transporter vermittelt

Die ungleichmäßige Genexpression in einer klonalen Zellkultur die zu Subpopulationen führt, wird gewöhnlich als Bistabilität bezeichnet. Dieses Phänomen wird in der Regel durch positive Rückkopplungsschleifen in Gen-Regelkreise verursacht (Dubnau and Losick, 2006). Im Falle des vorher beschriebenen strukturell ähnlichen *lac*-Operon in *E. coli* (Kapitel 1.3) wird nach Induktion der Genexpression mit Lactose ebenfalls ein bistabiles Verhalten (Ozbudak *et al.*, 2004) beobachtet. Der Repressor *LacI* wird von Lactose inaktiviert und dereprimiert Gene für den Lactoseabbau. Hier wird Bistabilität durch positive Rückkopplung über die verstärkte Bildung der Lactose-Permease *LacY*, die die Lactose-Aufnahme in die Zelle fördert und wiederum die Repression durch *LacI* hemmt, vermittelt.

Um zu testen, ob der skizzierte Fall auch auf die Beobachtung im Xyloseoperon von *B. megaterium* zutrifft, wurde die Rückkopplungsschleife von Transporter und Induktor in diesem System untersucht. Das Genexpressionsniveau des Repressors *xyIR* und Transporters *xyIT* ist mittels qRT-PCR sowohl in hochproduzierenden, als auch in niedrigproduzierenden Zellen näher analysiert worden. Induzierte Zellen wurden entsprechend ihres eGFP-Gehalts in zwei Fraktionen sortiert. Beide gesorteten Subpopulationen wurden auf A5-Festmedium mit 0,5 % Xylose ausplattiert. Dabei konnte gezeigt werden, dass diese Zellen immer noch in der Lage sind, große Mengen an eGFP herzustellen, was sowohl Mutationen als auch einen vollständigen Plasmidverlust ausschloß. In einem durchgeführten Plasmid Loss Versuch blieb das Plasmid über mehrere Stunden stabil bestehen. Die qRT-PCR-Analysen von sortierten Zellen aus beiden Subpopulationen zeigten die relativen Expressionsverhältnisse der beteiligten Gene. In hochproduzierenden Zellen wurde die Expression von *xyIR* etwa 150-fach und das Genprodukt eGFP ca. 500-fach induziert, wohingegen der Transporter *xyIT* unbeeinflusst blieb, obwohl er das letzte Gen des Operons *xyLABT* darstellt. Er scheint aus der XylR-vermittelten Regulation von *xyIA* und *xyIB* abgekoppelt zu werden, was durch die Existenz eines internen Transkriptionsterminators innerhalb des Operons stromaufwärts

von *xylT* (Schmiedel *et al.*, 1997) erklärt werden kann. Dennoch konnte selbst in *XylT*-Mutanten von *B. megaterium* die beobachtete Heterogenität (unveröffentlichte Daten) gezeigt werden. Somit scheint im Gegensatz zum *lac*-System, der beobachtete Phänotyp nicht über die Induktion der entsprechenden Zuckertransporter vermittelt zu werden. Die Beobachtung, dass viel Produkt an eGFP mit hohen Induktionsverhältnissen von *gfp* und *xylR* korreliert führte zu der Idee, dass die entstehende Heterogenität ein Gen-Dosis-Effekt ist, der höchstwahrscheinlich aus einer ungleichmäßigen Plasmidverteilung resultiert.

Um einen tieferen Einblick in die Dynamik der auftretenden Heterogenität einer Bakterienpopulation zu erhalten, wurde *XylR* mit mCherry fusioniert und Zeitraffer-Filme einer wachsenden Mikrokolonie mit dem Mikroskop aufgenommen. Dabei wurde sowohl die Produktion von eGFP, als auch von mCherry als Reaktion auf Xylose aufgezeichnet.

Am Anfang der exponentiellen Phasen konnte eine enorme Heterogenität verzeichnet werden, die sich zum Ende immer mehr in Richtung der hochproduzierenden Zellen verschob. Als Grund hierfür wurde sowohl die ungleichmäßige Verteilung des Xylose-Repressors, als auch des Plasmids in Betracht gezogen. Deshalb wurde ein Plasmid konstruiert, dem das *xylR*-Gen fehlte (pKMMBm5). Diese Mutante war keine echte Nullmutante, da das genomische *XylR* noch vorhanden war. Im Gegensatz zu der produzierten Menge des multicopy Plasmids konnte dies jedoch vernachlässigt werden. Als Resultat zeigte diese Mutante allerdings selbst im nicht induzierten Zustand weiterhin die beobachtete Heterogenität, so dass der Effekt nicht mit einer asymmetrischen Verteilung des Repressors erklärt werden konnte. Grundsätzlich konnte jedoch beobachtet werden, je höher die Konzentration an Xylose war, desto kleiner wurden die Zellen und um so schneller stellte sich die Heterogenität ein.

4.2 Populationsheterogenität wird durch die ungleichmäßige Plasmidverteilung belegt

Um konkrete molekulare Informationen über den Einfluss der Plasmidkopienzahl zu gewinnen, wurde das Plasmid visualisiert. Ähnlich wie bei einem fluoreszierenden

Repressor-Operator-System (FROS) für das *in vivo*-Tagging wurde als Fusionsprotein XylR-mCherry verwendet. Zusätzlich wurden Analysen mittels Fluoreszenz *in situ* Hybridisierung (FISH) zur direkten zellulären Lokalisation des Plasmids eingesetzt. Das Produktionsplasmid pSSBm85 zeigte nach der Induktion eine ungleichmäßige Verteilung der Plasmide mit dem Bestreben am Zellpol zu verweilen, während es im uninduzierten Zustand gleichverteilt blieb.

Die gewonnenen Daten legten nahe, dass selbst bei selektiven Bedingungen die beobachtete bimodale Verteilung kein Resultat von unterschiedlichen Gen-Regelkreisen war, sondern eine Frage der ungleichen Plasmidsegregation zwischen Tochterzellen.

Ein weiterer wichtiger Aspekt bei der Verteilung von multicopy Plasmiden ist ihre Größe, weshalb diesbezüglich das weitaus kleinere Vorgängermodell des eingesetzten Plasmids getestet wurde. Das weit verbreitete, winzige, multicopy Plasmid pBC16 stammt von Gram-positiven Bakterien ab und zeigt unter natürlichen Bedingungen eine Tetracyclinresistenz für Bazillen und einigen anderen Bodenbakterien (Polak und Novick, 1982). In diesem Fall sind mehrere kleine Foci vorhanden, die scheinbar zufällig über das ganze Cytoplasma verteilt sind, womit eindeutig die Größe eine entscheidende Rolle bei der Bildung von distinkten Foci spielt. Die generierten Ergebnisse gaben neue Einblicke in die Plasmid-Segregation von multicopy Plasmiden und haben erhebliche Folgen auf biotechnologische Proteinproduktionsprozesse. Aufgrund der Tatsache, dass die Plasmide im allgemeinen in geclusterten Foci auftreten und dadurch die freie Diffusion innerhalb von lebenden Zellen im allgemeinen begrenzt ist, wurde das Modell der zufälligen Diffusion in Frage gestellt. Es wurde postuliert, dass chromosomenkodierte Plasmide mindestens eine Kopienzahl dieser Cluster rekrutieren und die Partitionierung mittels einem bisher unbekannten Mechanismus vermitteln (Nordström and Gerdes, 2003 und Million-Weaver and Camps, 2014).

Das Plasmid pUB110, das verwandt zu dem eingesetzten Expressionsplasmid ist, verfügt über spezielle Anheftungsstellen (attachment site) an der Zellmembran (Tanaka and Sueoka, 1983 und Korn *et al.*, 1983). Allerdings belegten 3D-Analysen der konfokalen Laser Scanning Mikroskopie (CLSM) die Anordnung der Plasmide im zentralen Bereich der transaxialen Ebene, wobei die Membranbindestellen in den benutzten Expressionsplasmiden im Vergleich zum Ursprungsvektor pBC16 deutlich reduziert wurden.

Auf lange Sicht betrachtet wirkt sich die Anhäufung von Plasmiden durch die erhöhte metabolische Last negativ auf die Zellteilungsrate der Bakterien aus, da diese Produktionsrate die Zelle an ihre Kapazitätsgrenze bringen kann.

Eine andere Erklärung könnte mit der bakteriellen Zellalterung in Verbindung gebracht werden. Plasmide sammeln sich an den alten Polen an und die Vitalität der Mutterzellen wird mit zunehmender Anzahl der Generationen geringer (Stewart *et al.*, 2005 und Lindner *et al.*, 2008). Allerdings zeigt die Studie eine zufällige Auswahl der Zellenden, was damit diesen Effekt verringert. Vielleicht ist aber auch die aufwendige Energiebereitstellung für den Transport der Plasmide während der Zellteilung und damit die räumliche Trennung der stets neu replizierenden Plasmiden nicht möglich. Darüber hinaus könnte ein induziertes Plasmid in seiner ganzen Komplexität eine immense sterische Hinderung innerhalb einer organisierten Bakterienzelle darstellen.

Bislang ist nur wenig über die Verteilung von highcopy Plasmiden, insbesondere bei Gram-positiven Bakterien, bekannt (Wang *et al.*, 2004). Die Resultate zeigen eine polare Fixierung der Plasmide bei der Herstellung enormer Mengen an rekombinantem Protein in *B. megaterium*. Die molekularen Mechanismen hinter diesen Beobachtungen bleiben jedoch weitgehend unbekannt. Zukünftige Studien sind erforderlich, um Licht in die molekularen Grundlagen zu bringen.

5. Summary

Phenotypic heterogeneity during heterologous protein production in *Bacillus megaterium*

Within the last years *Bacillus megaterium* was systemically developed for the gram per liter production of recombinant proteins using the strong xylose-inducible promoter system P_{xylA} . The expression system is based on a multicopy plasmid containing the functional elements of the system – the gene encoding the repressor XylR and the promoter P_{xylA} . For a deeper understanding, P_{xylA} was fused to the coding sequence of the green fluorescent protein (GFP). In absence of xylose its expression is repressed by the xylose repressor XylR while in the presence of xylose the expression is derepressed. Although high yields of GFP were produced the culture showed a significant level of heterogeneity at the single cell level during the recombinant protein production process with up to 30 % of low-producing cells leading to three speculative working models or combinations of them – (1) an imbalanced repressor inducer equilibrium, (2) an unequal distribution of plasmids during cell division or (3) cell aging effects. In order to validate these models, new plasmids carrying different modifications were generated. For an estimation of the single cell XylR abundance and intracellular distribution the plasmid encoded *xylR* gene was translationally fused to *mCherry*. Further, the intracellular ratio of repressor molecules and DNA binding sites was modified by deleting this *xylR* gene and introducing additional XylR binding sites. Finally, alternative origins of replication were introduced to change and affect the plasmid copy number. All new plasmid strains were analyzed with respect to their overall GFP production behavior grown under different conditions. Single cell analyses were performed using time-lapse microscopy and flow cytometry to get an insight in the production dynamics of distinct cell lineages and subpopulation distributions. In addition, fluorescence *in situ* hybridization (FISH) experiments were performed to localize the plasmids in the cells. Derived results led to the hypothesis, that heterogeneous production behavior is probably a consequence of unequal plasmid distribution. Finally, the observations were incorporated into an extended mathematical model in order to reproduce the bistable production behavior via computer simulations.

6. Ausblick

Die gewonnenen Daten lassen den Schluss zu, dass die beobachtete Bimodalität keine Folge unterschiedlicher Genregulationen war, sondern vielmehr eine Frage der ungleichen Plasmidverteilung zwischen Tochterzellen, die selbst unter selektiven Bedingungen vorhanden war. Darüber hinaus wird die gegenwärtige Annahme über eine freie Diffusion von Plasmiden in Frage gestellt und ein Modell, das die ungleiche Verteilung von Plasmiden fördert, vorgeschlagen. Diese Resultate liefern neue Erkenntnisse über die heterogene Verteilung der multicopy Plasmiden und signalisieren erhebliche Folgen während des Produktionsprozesses rekombinanter Proteine (Fritzsche et al, 2012 sowie Neumeyer et al, 2013). Die Berücksichtigung der ungleichen Verteilung von Plasmiden könnte daraufhin ein wichtiges Kriterium für zukünftige biotechnologische Anwendungen sein.

Zur Weiterführung der Resultate wäre der nächste Schritt, pKMMBm6 genauer zu untersuchen, da dies das Plasmid mit der größten Heterogenität war. Zudem müsste auch der Einfluss auf die phänotypische Heterogenität von eGFP mittels pKMMBm7 untersucht werden. Dieses Plasmid trägt ausschließlich mCherry als Fusionsprotein, so dass es sich sehr gut für Time-Lapse Mikroskopie eignet. Ein weiterer spannender Aspekt wäre die Zellalterung und das Phänomen des eGFP-Abbaus am Ende der stationären Phase im halbdefinierten A5-Medium. Zudem könnte der Einfluss der Induktion exakter anhand neuer Flussmethoden (Mikrofluidik) am Mikroskop untersucht werden. Andere Ansätze zur Bestimmung der PCN könnten verfolgt werden und eventuell genauere Daten liefern. Zu guter Letzt würde ein Microarray Einblicke in die Vorergebnisse der Hormesis von subletalen Antibiotikakonzentrationen geben.

7. Referenzen

Acar, M., Mettetal, J.T., van Oudenaarden, A. (2008). Stochastic switching as a survival strategy in fluctuating environments. *Nat Genet* **40**, 471-475.

Ackermann, M., Stecher, B., Freed, N.E., Songhet, P., Hardt, W.-D. and Doebeli, M. (2008). Self-destructive cooperation mediated by phenotypic noise. *Nature* **454**, 987-990.

Belitsky, B.R., Kim, H.J., Sonenshein, A.L. (2004). CcpA-dependent regulation of *Bacillus subtilis* glutamate dehydrogenase gene expression. *J Bacteriol* **186**, 3392–3398.

Bernhard, K., Schrempf, H., and Goebel, W. (1978). Bacteriocin and antibiotic resistance plasmids in *Bacillus cereus* and *Bacillus subtilis*. *J Bacteriol* **133**, 897–903.

Biedendieck, R., Yang, Y., Deckwer, W.-D., Marco Malten, and Jahn, D. (2007). Plasmid system for the intracellular production and purification of affinity-tagged proteins in *Bacillus megaterium*. *Biotechnol Bioeng* **96**, 525–537.

Biedendieck, R., Bunk, B., Fürch, T., Franco-Lara, E., Jahn, M., and Jahn, D. (2010a). Systems biology of recombinant protein production in *Bacillus megaterium*. *Adv Biochem Eng Biotechnol* **120**, 133–161.

Biedendieck, R., Malten, M., Barg, H., Bunk, B., Martens, J.-H., Deery, E., Leech, H., Warren, M. J., and Jahn, D. (2010b). Metabolic engineering of cobalamin (vitamin b12) production in *Bacillus megaterium*. *Microb Biotechnol* **3**(1), 24–37.

Bray, C., Spidlen, J. and Brinkman, R.R. (2012). FCS 3.1 Implementation Guidance. *Cytometry A* **81**, 523-526.

Bron, S. and Luxen, E. (1985). Segregational instability of pUB110-derived recombinant plasmids in *Bacillus subtilis*. *Plasmid* **14**, 235-244.

Bunk, B., Schulz, A., Stammen, S., Münch, R., Warren, M.J., Rohde, M., Jahn, D., and Biedendieck, R. (2010). A short story about a big magic bug. *Bioeng Bugs* **1**, 85–91.

Bustin S.A., Benes V., Garson J.A., Hellemans J., Huggett J., Kubista M., Mueller R., Nolan T., Pfaffl M.W., Shipley G.L., Vandesompele J. and Wittwer C.T. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry* **55**, 611–622.

Butler, Y.X., Abhayawardhane Y. , Stewart G.C. (1993). Amplification of the *Bacillus subtilis maf* gene results in arrested septum formation. *J Bacteriol* **175**, 3139–3145.

Chopra I. and Roberts M. (2001). Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiol Mol Biol Rev* **65**, 232–260.

Cohn, M. and Horbita, K. (1959) Analysis of the differentiation and of the heterogeneity within a population of *Escherichia coli* undergoing induced beta-galactosidase synthesis. *J Bacteriol* **78**, 613–623.

De Bary, A. (1884). Vergleichende Morphologie und Biologie der Pilze, Mycetozoen und Bakterien. Wilhelm Engelmann, Leipzig, Germany.

De Jong, I.G., Beilharz, K., Kuipers, O.P. and Veening, J.-W. (2011). Live cell imaging of *Bacillus subtilis* and *Streptococcus pneumoniae* using automated time-lapse microscopy. *J Vis Exp* **53**, 3145.

Dernburg, A.F. and Sedat, J.W. (1998). Mapping three-dimensional chromosome architecture *in situ*. *Methods Cell Biol* **53**, 187–233.

Dong, H., and Zhang, D. (2014). Current development in genetic engineering

strategies of *Bacillus species*. Microb Cell Fact **13**, 63.

Dubnau, D., and Losick, R. (2006). Bistability in bacteria. Mol Microbiol **61**, 564–572.

Ehrenreich, A. (2006). DNA microarray technology for the microbiologist: an overview. Appl Microbiol Biotechnol **73**, 255–273.

Ellis, B., Gentleman, R., Hahne, F., Meur, N.L. and Sarkar, D. flowViz: Visualization for flow cytometry. R package version 1.32.0.

Ellis, B., Haaland, P., Hahne, F., Meur, N.L., Gopalakrishnan, N., Spidlen, J. and Jiang M. flowCore: flowCore: Basic structures for flow cytometry data. R package version 1.34.10.

Eppinger, M., Bunk, B., Johns, M.A., Edirisinghe, J.N., Kutumbaka, K.K., Koenig, S.S.K., Creasy, H.H., Rosovitz, M.J., Riley, D.R., Daugherty, S., Martin M., Elbourne L.D., Paulsen I., Biedendieck R., Braun C., Grayburn S., Dhingra S., Lukyanchuk V., Ball B., Ul-Qamar R., Seibel J., Bremer E., Jahn D., Ravel J. and Vary P.S. (2011). Genome Sequences of the Biotechnologically Important *Bacillus megaterium* Strains QM B1551 and DSM 319. J Bacteriol **193**, 4199–4213.

Fleige, S., and Pfaffl, M.W. (2006). RNA integrity and the effect on the real-time qRT-PCR performance . Molecular Aspects of Medicine **27**,126–139.

Fritzsche, F.S.O., Dusny, C., Frick, O., and Schmid, A. (2012). Single-cell analysis in biotechnology, systems biology, and biocatalysis. Annu Rev Chem Biomol Eng **3**, 129–155.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A.J., Sawitzki, G., Smith,

C., Smyth, G., Tierney, L., Yang, J.YH and Zhang J. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**, R80.

Gordon, G.S., Sitnikov, D., Webb, C.D., Teleman, A., Straight, A., Losick, R., Murray, A.W., and Wright, A. (1997). Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. *Cell* **90**, 1113–1121.

Hahne, F., LeMeur, N., Brinkman, R.R., Ellis, B., Haaland, P., Sarkar, D., Spidlen, J., Strain, E., and Gentleman, R. (2009). flowCore: a Bioconductor package for high throughput flow cytometry. *BMC Bioinformatics* **10**, 106.

Heyndrickx, M. (2011). Dispersal of aerobic endospore-forming bacteria from soil and agricultural activities to food and feed. In *Aerobic, Endospore-forming Soil Bacteria* ed. Logan, N.A. and De Vos, P. pp. 135–156 Berlin: Springer-Verlag.

Hruz, T., Wyss, M., Docquier, M., Pfaffl, M.W., Masanetz, S., Borghi, L., Verbrugge, P., Kalaydjieva, L., Bleuler, S., Laule, O., Descombes, P., Gruissem, W. and Zimmermann, P. (2011). RefGenes: identification of reliable and condition specific reference genes for RT-qPCR data normalization. *BMC Genomics* **12**, 156.

Huber, W., Carey, J.V, Gentleman, R., Anders, S., Carlson, M., Carvalho, S.B, Bravo, C.H, Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, D.K, Irizarry, A.R, Lawrence, M., Love, I. M., MacDonald, J., Obenchain, V., Oleś, K.A, Pagés, H., Reyes, A., Shannon, P., Smyth, K.G, Tenenbaum, D., Waldron, L., Morgan, M. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods*, **12**, 115–121.

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Mol Microbiol* **3**, 318–356.

Jensen, R.B., and Shapiro, L. (1999). The *Caulobacter crescentus* *smc* gene is required

for cell cycle progression and chromosome segregation. *Proc Natl Acad Sci U A* **96**, 10661–10666.

Kaltwasser, M., Wiegert, T. and Schumann, W. (2002). Construction and application of epitope and green fluorescent protein-tagging integration vectors for *Bacillus subtilis*. *Appl Env Microbiol* **68**, 2624–2628.

Klein, J., Leupold, S., Biegler, I., Biedendieck, R., Münch, R., and Jahn, D. (2012). TLM-Tracker: software for cell segmentation, tracking and lineage analysis in time-lapse microscopy movies. *Bioinformatics* **28**, 2276–2277.

Kohl, M. (2007). SLqPCR: Functions for analysis of real-time quantitative PCR data at SIRSLab GmbH. R package, SIRS-Lab GmbH, Jena.

Korn, R., Winston, S., Tanaka, T., and Sueoka, N. (1983). Specific in vitro binding of a plasmid to a membrane fraction of *Bacillus subtilis*. *Proc Natl Acad Sci U A* **80**, 574–578.

Kussell, E. and Leibler, S. (2005). Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309**, 2075–2078.

Laishley, E.J., Bernlohr, R.W. (1968). Regulation of arginine and proline catabolism in *Bacillus licheniformis*. *J Bacteriol* **96**, 322–329.

Landgraf, D., Okumus, B., Chien, P., Baker, T.A., Paulsson, J. (2012). Segregation of molecules at cell division reveals native protein localization. *Nat Methods* **9**, 480–482.

Lefever, S., Hellemans, J., Pattyn, F., Przybylski, D.R., Taylor, C., Geurts, R., Untergasser, A., and Vandesompele, J. (2009). RDML: structured language and reporting guidelines for real-time quantitative PCR data. *Nucleic Acids Res* **37**, 2065–2069.

- Lindner, A.B., Madden, R., Demarez, A., Stewart, E.J., and Taddei, F.** (2008). Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc Natl Acad Sci U A* **105**, 3076–3081.
- Liu, L., Liu, Y., Shin, H.-D., Chen, R.R., Wang, N.S., Li, J., Du, G., and Chen, J.** (2013). Developing *Bacillus spp.* as a cell factory for production of microbial enzymes and industrially important biochemicals in the context of systems and synthetic biology. *Appl Microbiol Biotechnol* **97**, 6113–6127.
- Livak, K.J. and Schmittgen, T.D.** (2001). Analysis of Relative Gene Expression Data Using Real Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* **25**, 402-8.
- Locke, J.C.W. and Elowitz, M.B.** (2009). Using movies to analyse gene circuit dynamics in single cells. *Nat Rev Microbiol* **7**, 383–392.
- Malten, M., Hollmann, R., Deckwer, W.-D., and Jahn, D.** (2005). Production and secretion of recombinant *Leuconostoc mesenteroides* dextranucrase DsrS in *Bacillus megaterium*. *Biotechnol Bioeng* **89**, 206–218.
- Malten, M., Biedendieck, R., Gamer, M., Ann-Christin Drews, Stammen, S., Buchholz, K., Dijkhuizen, L., and Jahn, D.** (2006). A *Bacillus megaterium* plasmid system for the production, export, and one-step purification of affinity-tagged heterologous levansucrase from growth medium. *Appl Env. Microbiol* **72**, 1677–1679.
- Mascarenhas, J., Weber, M.H., and Graumann, P.L.** (2001). Specific polar localization of ribosomes in *Bacillus subtilis* depends on active transcription. *EMBO Rep* **2**, 685–689.
- Meinhardt, F., Stahl, U., and Ebeling, W.** (1989). Highly efficient expression of homologous and heterologous genes in *Bacillus megaterium*. *Appl Microbiol Biotechnol* **30**, 343–350.

Million-Weaver, S., and Camps, M. (2014). Mechanisms of plasmid segregation: have multicopy plasmids been overlooked? *Plasmid* **75**, 27–36.

Muench, K., Müller, J., Wienecke, S., Bergmann, S., Heyber, s., Biedendieck, R., Muench, R. and Jahn D. (2015). Polar fixation of plasmids during recombinant protein production in *Bacillus megaterium* results in population heterogeneity. *Environ Appl Microbiol* **81**, 5976–5986.

Neumeyer, A., Hübschmann, T., Müller, S., and Frunzke, J. (2013). Monitoring of population dynamics of *Corynebacterium glutamicum* by multiparameter flow cytometry. *Microb Biotechnol* **6**, 157–167.

Nordström, K., and Gerdes, K. (2003). Clustering versus random segregation of plasmids lacking a partitioning function: a plasmid paradox? *Plasmid* **50**, 95–101.

Novick, A. and Weiner, M. (1957). Enzyme induction as an all-or-none phenomenon. *Proc Natl Acad Sci U A* **43**, 553-566.

Ozbudak, E.M., Thattai, M., Lim, H.N., Shraiman, B.I., and Oudenaarden, A.V. (2004). Multistability in the lactose utilization network of *Escherichia coli*. *Nature* **427**, 737–740.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**, e45.

Pogliano, J., Ho, T.Q., Zhong, Z., and Helinski, D.R. (2001). Multicopy plasmids are clustered and localized in *Escherichia coli*. *Proc Natl Acad Sci U A* **98**, 4486–4491.

Polak, J., and Novick, R.P. (1982). Closely related plasmids from *Staphylococcus aureus* and soil bacilli. *Plasmid* **7**, 152–162.

Popov, M., Petrov, S., Nacheva, G., Ivanov, I., and Reichl, U. (2011). Effects of a

recombinant gene expression on ColE1-like plasmid segregation in *Escherichia coli*. BMC Biotechnol **11**, 18.

Reyes-Lamothe, R., Tran, T., Meas, D., Lee, L., Li, A.M., Sherratt, D.J., and Tolmasky, M.E. (2013). High-copy bacterial plasmids diffuse in the nucleoid-free space, replicate stochastically and are randomly partitioned at cell division. Nucleic Acids Res **42**, 1042-1051.

Ritz, C., and Spiess, A.-N. (2008). qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. Bioinformatics **24**, 1549–1551.

Rutledge, R.G., Stewart, D. (2008). Critical evaluation of methods used to determine amplification efficiency refutes the exponential character of real-time PCR. BMC Molecular Biology **9**, 96.

Rygus, T., and Hillen, W. (1991). Inducible high-level expression of heterologous genes in *Bacillus megaterium* using the regulatory elements of the xylose-utilization operon. Appl Microbiol Biotechnol **35**, 594–599.

Rygus, T., and Hillen, W. (1992). Catabolite repression of the xyl operon in *Bacillus megaterium*. J Bacteriol **174**, 3049–3055.

Rygus, T., Scheler, A., Allmansberger, R., and Hillen, W. (1991). Molecular cloning, structure, promoters and regulatory elements for transcription of the *Bacillus megaterium* encoded regulon for xylose utilization. Arch Microbiol **155**, 535–542.

Sánchez-Romero, M.-A., Lee, D.J., Sánchez-Morán, E. and Busby, S.J.W. (2012). Location and dynamics of an active promoter in *Escherichia coli* K-12. Biochem J **441**, 481–485.

Sarkar, D., Le Meur, N., and Gentleman, R. (2008). Using flowViz to visualize flow

cytometry data. *Bioinformatics* **24**, 878–879.

Schmiedel, D., Kintrup, M., Küster, E., and Hillen, W. (1997). Regulation of expression, genetic organization and substrate specificity of xylose uptake in *Bacillus megaterium*. *Mol Microbiol* **23**, 1053–1062.

Scholz, O., Thiel, A., Hillen, W., Niederweis, M. 2000. Quantitative analysis of gene expression with an improved green fluorescent protein. *Eur J Biochem* **267**, 1565–1570.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M. and Ragg T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* **7**, 3.

Shi, L., Guenther, S., Huebschmann, T., Wick, L.Y., Harms, H., Mueller, S. (2007). Limits of Propidium Iodide as a Cell Viability Indicator for Environmental Bacteria. *Cytometry Part A* **71A**, 592–598.

Škulj, M., Okršlar, V., Jalen, Š., Jevševar, S., Slanc, P., Štrukelj, B., and Menart, V. (2008). Improved determination of plasmid copy number using quantitative real-time PCR for monitoring fermentation processes. *Microbial cell factories* **7**(1), 6.

Sojka, L., Fucík, V., Krásný, L., Barvík, I. and Jonák, J. (2007). YbxF, a protein associated with exponential-phase ribosomes in *Bacillus subtilis*. *J Bacteriol* **189**, 4809–4814.

Solopova, A., van Gestel, J., Weissing, F.J., Bachmann, H., Teusink, B., Kok, J. and Kuipers O.P. (2014). Bet-hedging during bacterial diauxic shift. *Proc Natl Acad Sci U S A* **111**, 7427–7432.

Soufo, H.J.D., and Graumann, P.L. (2010). *Bacillus subtilis* MreB paralogues have different filament architectures and lead to shape remodelling of a heterologous cell

system. Mol Microbiol **78**, 1145–1158.

Spiess, A.-N., Feig, C., and Ritz, C. (2008). Highly accurate sigmoidal fitting of real-time PCR data by introducing a parameter for asymmetry. BMC Bioinformatics **9**, 221.

Stammen, S., Müller, B.K., Korneli, C., Rebekka Biedendieck, Gamer, M., Franco-Lara, E., and Dieter Jahn (2010). High-yield intra- and extracellular protein production using *Bacillus megaterium*. Appl Env. Microbiol **76**, 4037–4046.

Stewart, E.J., Madden, R., Paul, G., and Taddei, F. (2005). Aging and death in an organism that reproduces by morphologically symmetric division. PLoS Biol **3**, e45.

Straight, A.F., Belmont, A.S., Robinett, C.C., and Murray, A.W. (1996). GFP tagging of budding yeast chromosomes reveals that protein-protein interactions can mediate sister chromatid cohesion. Curr Biol **6**, 1599– 1608.

Summers, D.K. (1996). The Biology of Plasmids (Blackwell Science Ltd), 68 – 82.

Tanaka, T., and Sueoka, N. (1983). Site-specific in vitro binding of plasmid pUB110 to *Bacillus subtilis* membrane fraction. J Bacteriol **154**, 1184–1194.

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol **72**, 211–222.

Thellin, O., Zorzi, W., Lakaye, B., De Borman, B., Coumans, B., Hennen, G., Grisar, T., Igout, A. and Heinen, E. (1999): Housekeeping genes as internal standards: use and limits, J. Biotechnol **75**, 291–295.

Tichopad, A., Dilger, M., Schwarz, G., and Pfaffl, M.W. (2003). Standardized determination of real-time PCR efficiency from a single reaction set-up. Nucleic Acids Res **31**, e122.

Tricarico, C., Pinzani, P., Bianchi, S., Paglierani, M., Distante, V., Pazzagli, M., Bustin, S.A. and Orlando, C. (2002): Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. *Anal Biochem* **309**, 293–300.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., and Leunissen, J.A.M. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res* **35**, W71–W74.

Vandesompele, J., Preter, K.D., Pattyn, F., Bruce Poppe, Roy, N.V., Paepe, A.D., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**, RESEARCH0034.

Vary, P. S. (1994). Prime time for *Bacillus megaterium*. *Microbiology*, **140** (Pt 5), 1001–1013.

Vary, P.S., Biedendieck, R., Fuerch, T., Meinhardt, F., Rohde, M., Deckwer, W.-D., and Jahn, D. (2007). *Bacillus megaterium* – from simple soil bacterium to industrial protein production host. *Appl Microbiol Biotechnol* **76**, 957–967.

Voigt B., Schroeter R., Schweder T., Jürgen B., Albrecht D., van Dijl J.M., Maurer K.H., Hecker M. (2014). A proteomic view of cell physiology of the industrial workhorse *Bacillus licheniformis*. *J Biotechnol.* **191**, 139-149.

Wang, J.D., Rokop, M.E., Barker, M.M., Hanson, N.R., and Grossman, A.D. (2004). Multicopy plasmids affect replisome positioning in *Bacillus subtilis*. *J Bacteriol* **186**, 7084–7090.

Yang, Y.H. and Thorne N.P. (2003). Normalization for two-color cDNA microarray data. *The Institute of Mathematical Statistics Lecture Notes - Monograph Series* **40**, 403-418.

Young, J.W., Locke, J.C.W., Altinok, A., Rosenfeld, N., Bacarian, T., Swain, P.S., Mjolsness, E., and Elowitz, M.B. (2012). Measuring single-cell gene expression dynamics in bacteria using fluorescence time-lapse microscopy. *Nat Protoc* 7, 80–88.

8. Anhang

8.1 Filme

- Film 1:** Zeitraffer-Mikroskopie-Film von *B. megaterium*, der das eGFP enthaltende Plasmid pSSBm85 enthält. Dieser Film veranschaulichte die Wachstumsdauer über einen Zeitraum von 5,4 Std. Die produzierenden Mikrokolonien zeigten dabei eine steigende Menge an eGFP. Als Folge der Zellteilung ergaben sich Zellen, die sowohl eine hohe, als auch eine niedrige Produktionsrate erzeugten. Dieses Phänomen resultierte in einer Populationsheterogenität.
- Film 2:** Zeitraffer-Film einer Zweikanal-Fluoreszenzmikroskopie über einen Zeitraum von 7,8 Std. Dieser Film beinhaltete eine wachsende Mikrokolonie von *B. megaterium*, mit dem Plasmid pKMMBm2. Dieser Vektor produzierte eGFP (grüne Fluoreszenz) und XylR-mCherry (rote Fluoreszenz). Parallel konnte die Fluoreszenzverteilung visualisiert werden.
- Film 3:** Bei diesem Zweikanal-Zeitraffer-Mikroskopie-Film handelte es sich abermals um eine wachsende Mikrokolonie von *B. megaterium*, mit dem pKMMBm2 Plasmid. Er zeigte einen Zeitraum von 2 Std und beinhaltete eine Zelle, die XylR-mCherry anhäufte und dabei die eGFP-Produktion eingestellt.
- Film 4:** Zeitraffer-Mikroskopie-Film einer Mikrokolonie von *B. megaterium*, der einen Zeitraum von 4,2 Std zeigte. Dieser Stamm enthielt das Plasmid pKMMBm5. Dabei handelte es sich um ein Derivat von pSSBm85 – ein Vektor, der kein *xylR* mehr beinhaltete und somit ohne Induktion bereits eGFP produzierte. Die Produktionsheterogenität blieb in Abwesenheit von dem plasmidkodierten *xylR*-Gen bestehen.
- Film 5:** Dieser Film zeigte den selben Stamm, wie in Film 4 beschrieben, nur dass dieser bereits von Anfang an induziert wurde. Auch hier konnte eine Heterogenität beobachtet werden.

8.2 R-Skripte zur Datenauswertung

Prinzipiell mussten alle Parameter individuell diesem Grundgerüst angepaßt werden.

8.2.1 qRT-PCR

8.2.1.1 R-Skript zur Berechnung der Effizienz und des C_q -Wertes, sowie zum Erstellen der Boxplots

```
#-----  
# qpcrrun.R  
#  
# Analysiert und plottet 2x48 qPCR Diagramme mit Cp und Effizienz  
#  
# Angepasst für BIO-RAD Real Time PCR CFX96  
#  
# Daten Vorverarbeitung:  
# Export der Rohdaten als Excel File über die BIO-RAD CFXManager Software  
# Öffnen der Datei mit Open/Libre Office und speichern als csv (comma separated  
# values)  
#  
# Voraussetzungen:  
# R Version > 2.13          (http://cran.r-project.org/)  
# qpcR package 1.3-6       (http://www.dr-spiess.de/ oder  
#                           http://cran.r-project.org/src/contrib/Archive/qpcR/)  
# to do:  
# - Analyse von Daten verteilt auf mehreren Platten  
#  
# Letzte Änderung: 2013-01-11  
#-----  
  
# qpcR Bibliothek laden:  
library(qpcR)  
  
# Eingabeparameter  
  
# Projektname:  
# Hinweise zum Projektnamen:  
# Dateiname zum Einlesen der Rohdaten: <project>.csv  
# Dateiname zum Schreiben der berechneten qPCR Parameter: <project>_analyzed.csv  
# Dateiname zum Schreiben der Grafiken als PDF: <project>.pdf  
project <- "kmm64"  
# Art der Auswertung  
#job <- "view" # Berechnung aller qPCR parameters und plotten  
job <- "all"  # zusätzlich Berechnung der Ratios  
# Aufbau der CSV Datei
```

```

csvdec <- "," # Dezimaltrennzeichen
csvsep <- ";" # Spaltentrennzeichen

# Ausgabe in Grafikdatei (TRUE/FALSE)
# Alternativ Ausgabe am Bildschirm
graphics <- FALSE
graphics <- TRUE

# Regressionsmodell
# benutztes sigmoides Modell zur nicht-linearen Regression der qPCR
nlmodel <- 15 # alternativ: 14, 14, 15, 16, 17, b3, b4, b5, b6, b7

# Art der Fehleranalyse
# erroranalysis <- "sims" # Monte Carlo simulation
erroranalysis <- "perms" # Permutation approach
# erroranalysis <- "props" # Error propagation

# Layout der 96-Well Platte
# Alle Wells, die nicht analysiert werden sollen, müssen weggelassen werden!
# (Bei unserer Real-Time-PCR-Maschine ist keine passive Referenz notwendig, beim
Auswerten
# mehrerer wells csv-Dateien und platelayout erweitern mit I1-I12, J1-J12 etc.)
# platelayout <-
#c("A1", "A2", "A3", "A4", "A5", "A6", "A7", "A8", "A9", "A10", "A11", "A12",
# "B1", "B2", "B3", "B4", "B5", "B6", "B7", "B8", "B9", "B10", "B11", "B12",
# "C1", "C2", "C3", "C4", "C5", "C6", "C7", "C8", "C9", "C10", "C11", "C12",
# "D1", "D2", "D3", "D4", "D5", "D6", "D7", "D8", "D9", "D10", "D11", "D12",
# "E1", "E2", "E3", "E4", "E5", "E6", "E7", "E8", "E9", "E10", "E11", "E12",
# "F1", "F2", "F3", "F4", "F5", "F6", "F7", "F8", "F9", "F10", "F11", "F12",
# "G1", "G2", "G3", "G4", "G5", "G6", "G7", "G8", "G9", "G10", "G11", "G12",
# "H1", "H2", "H3", "H4", "H5", "H6", "H7", "H8", "H9", "H10", "H11", "H12");

# platelayout <-
#c("A1", "A2", "A3", "A4", "A5", "A6", "A7", "A8", "A9", "A10", "A11", "A12",
# "B1", "B2", "B3", "B4", "B5", "B6", "B7", "B8", "B9", "B10", "B11", "B12",
# "C1", "C2", "C3", "C4", "C5", "C6", "C7", "C8", "C9", "C10", "C11", "C12",
# "D1", "D2", "D3", "D4", "D5", "D6", "D7", "D8", "D9", "D10", "D11", "D12",
# "E1", "E2", "E3", "E4", "E5", "E6", "E7", "E8",
# "F1", "F2", "F3", "F4", "F5", "F6", "F7", "F8",
# "G1", "G2", "G3", "G4", "G5", "G6", "G7", "G8",
# "H1", "H2", "H3", "H4", "H5", "H6", "H7", "H8");

# Gruppierung der einzelnen Proben (nur bei Ratio Berechnung)
# r: reference gene, g: gene of interest; s: target sample, c: control sample
# Beispiel:
# "g1s1": gene-of-interest #1 in target sample #1
# "g1c1": gene-of-interest #1 in control sample #1
# "r1s1": reference gene #1 in target sample #1

```

```
# "r1c1": reference gene #1 in control sample #1
plategroups <-
c("r1c1", "r1c1", "r2c1", "r2c1", "r3c1", "r3c1", "r4c1", "r4c1", "g1c1", "g1c1", "g2c1",
  "g2c1",
  "r1s1", "r1s1", "r2s1", "r2s1", "r3s1", "r3s1", "r4s1", "r4s1", "g1s1", "g1s1", "g2s1",
  "g2s1",
  "r1s2", "r1s2", "r2s2", "r2s2", "r3s2", "r3s2", "r4s2", "r4s2", "g1s2", "g1s2", "g2s2",
  "g2s2",
  "r1s3", "r1s3", "r2s3", "r2s3", "r3s3", "r3s3", "r4s3", "r4s3", "g1s3", "g1s3", "g2s3",
  "g2s3",
  "g3c1", "g3c1", "g4c1", "g4c1", "g5c1", "g5c1", "r5c1", "r5c1",
  "g3s1", "g3s1", "g4s1", "g4s1", "g5s1", "g5s1", "r5s1", "r5s1",
  "g3s2", "g3s2", "g4s2", "g4s2", "g5s2", "g5s2", "r5s2", "r5s2",
  "g3s3", "g3s3", "g4s3", "g4s3", "g5s3", "g5s3", "r5s3", "r5s3");

# Gen-Namen (Reihenfolge definiert in plategroups: g1, g2, ..., gn), die Zuordnung der
# Variablen r in den gennames ist nicht gleichzusetzen mit der Zuordnung der plategroups
geneorder<-
c("xylR_1", "xylR_2", "xylR_3", "xylT_1", "xylT_2", "xylT_3", "gfp_1", "gfp_2", "gfp_3",
  "xylA_1", "xylA_2", "xylA_3", "xylB_1", "xylB_2", "xylB_3")

genenames <- data.frame("xylR_1" = "g1c1:g1s1:r1c1:r1s1",
                        "xylT_1" = "g2c1:g2s1:r1c1:r1s1",
                        "gfp_1" = "g3c1:g3s1:r1c1:r1s1",
                        "xylA_1" = "g4c1:g4s1:r1c1:r1s1",
                        "xylB_1" = "g5c1:g5s1:r1c1:r1s1",
                        "xylR_2" = "g1c1:g1s2:r1c1:r1s2",
                        "xylT_2" = "g2c1:g2s2:r1c1:r1s2",
                        "gfp_2" = "g3c1:g3s2:r1c1:r1s2",
                        "xylA_2" = "g4c1:g4s2:r1c1:r1s2",
                        "xylB_2" = "g5c1:g5s2:r1c1:r1s2",
                        "xylR_3" = "g1c1:g1s3:r1c1:r1s3",
                        "xylT_3" = "g2c1:g2s3:r1c1:r1s3",
                        "gfp_3" = "g3c1:g3s3:r1c1:r1s3",
                        "xylA_3" = "g4c1:g4s3:r1c1:r1s3",
                        "xylB_3" = "g5c1:g5s3:r1c1:r1s3")

#-----
# ab hier Berechnung...

# Dateiname mit Rohdaten
csvfile <- paste(project, ".csv", sep="")

# Dateiname für PDF Datei
pdffile <- paste(project, "_genes", ".pdf", sep="")
# Dateiname für ausgewertete qPCR Daten
qpcrfile <- paste(project, "_analyzed", ".csv", sep="")
```

```
# Grafikausgabe in Datei, falls ausgewählt
if (graphics) {pdf(pdffile,width=8,height=5,pointsize=4,paper="special")}

# CSV Datei öffnen:
rawdata <- read.csv2(file = csvfile, sep = csvsep, dec = csvdec);
# rawdata <- read.csv2(file = "clipboard", sep = csvsep, dec = csvdec); # aus der
Zwischenablage

# erste Spalte löschen da immer leer nach xls -> csv Export mit Libre Office
rawdata <- rawdata[,-c(1)];

# die "Cycle" Spalte in "Cycles" umbenennen (für die pcrbatch Funktion erforderlich)
names(rawdata)[1] <- "Cycles";

# nicht ausgewählte Wells weglassen
rawdata2 <- rawdata[c(1, match(platelayout, names(rawdata)))]

# die Daten können ggf. mit View betrachtet werden
# Hinweis: die 1. Spalte enthält die Zyklenanzahl
# ab der 2. Spalte kommen die Daten der einzelnen Wells (A1..H12)
# View(rawdata)

# Berechnung aller qPCR parameters
# qpcrdata <- pcrbatch(rawdata2, model = nlmodel, do.mak = FALSE, plot = FALSE);
qpcrdata <- pcrbatch(rawdata2, model = nlmodel, plot = FALSE);
# Daten in Datei schreiben
# Vorsicht: alle Daten sind Characters in Anführungszeichen, d.h. dec hat da keinen
Einfluss
write.table(qpcrdata, file = qpcrfile, append = FALSE, sep = csvsep, dec = csvdec);

# Einzelberechnungen (nur zur Info):
# mod_15 <- pcrfit(rawdata, 1, i, 15)
# mod_mak3 <- pcrfit(rawdata, 1, i, mak3)
# mod_chag <- pcrfit(rawdata, 1, i, chag)
# plot(mod_15)
# efficiency(mod_15, type = "cpD2", plot = FALSE)$eff
# expfit(mod_15, plot = FALSE)$eff
# sliwin(mod_15, plot=TRUE)$eff
# Diagramme plotten

startnumber <- 2; # erste Datenspalte
for(p in 1:2) { # in 2x48 wells aufteilen
  # Layout 8x6 Grafiken (insgesamt 48), Ränder anpassen
  # halbe 96-well Platte
  par(mfrow=c(6,8),mar=c(2,2,1.5,0), omi=c(0,0,0,0.1));

  plotnumber <- length(qpcrdata)
```

```

if ((plotnumber > 49) && (p == 1) ) plotnumber <- 49;

if (startnumber <= plotnumber){
  for(i in startnumber:plotnumber) {

    # Daten mit Regressionskurve plotten
    mod_qpcr <- pcrfit(rawdata2, 1, i, nlmodel);
    plot(mod_qpcr);

    # die wichtigsten Kanndaten aus der Datenmatrix holen
    cpD2      <- as.numeric(qpcrdata[which(qpcrdata[,1]=="sig.cpD2"),i]);
    fluo      <- as.numeric(qpcrdata[which(qpcrdata[,1]=="sig.fluo"),i]);
    sig.eff   <- as.numeric(qpcrdata[which(qpcrdata[,1]=="sig.eff"),i]);
    exp.eff   <- as.numeric(qpcrdata[which(qpcrdata[,1]=="exp.eff"),i]);
    sli.eff   <- as.numeric(qpcrdata[which(qpcrdata[,1]=="sli.eff"),i]);

    # gestrichelte Linie des berechneten Cp Werts
    abline(h=fluo, lty=2) # horizontale Baseline
    abline(v=cpD2, lty=2) # vertical

    # Titel: Well + Cq Wert
    column_name <- attributes(rawdata2)$names[i];
    title(main= bquote(paste(.(column_name), " ", C[q], ": ", .(cpD2))));

    # Effizienzwerte (sig, exp und sli) neben Y-Achse ausgeben:
    maxY <- max(rawdata2[,i])
    text(10,maxY/4*3,paste("sig.eff: ",round(sig.eff,2)))
    text(10,maxY/4*2,paste("exp.eff: ",round(exp.eff,2)))
    text(10,maxY/4*1,paste("sli.eff: ",round(sli.eff,2)))
  } # for i
} # if (startnumber <= plotnumber)

startnumber <- plotnumber + 1

} # for p

if (job == "all") { # Berechnung der Ratios, falls job == "all"

  ratios <- ratiobatch(qpcrdata, group = plategroups, plot = TRUE, which.eff = "sig",
refmean = T)
  eatype <- c("sims", "perms", "props")
  E <- match(erroranalysis, eatype)
  ratiodef <- data.frame(ratios$rawData[[1]][,E])
  if (ncol(genenames) > 1) {
    for (g in 2:ncol(genenames)) {
      ratiodef <- cbind(ratiodef, ratios$rawData[[g]][,E])
    } # for g
  } # if
}

```



```

colnames(ratiodf) <- names(ratios$resList)

par(mfrow=c(1,1), mar=c(8,8,8,8)) # neue Seite beginnen
# plot.new()

boxplot(ratiodf[as.vector(t(genenames[geneorder]))], names = geneorder, outline =
FALSE, log = "y", col="lightgreen")

} # if job == "all"

if (graphics) { dev.off() } # gegebenenfalls Grafikdatei schreiben

# ## plot results
# if (plot && length(outLIST) < 50) {
#   DIM <- ceiling(sqrt(length(outLIST)))
#   par(mfrow = c(DIM, DIM + 1))
#   par(mar = c(1, 4, 2, 2))
#   for (i in 1:length(outLIST)) {
#     outDATA[[i]][outDATA[[i]] <= 0] <- NA
#     YLIM <- c(min(outDATA[[i]], na.rm = TRUE), max(outDATA[[i]], na.rm =
TRUE))
#     boxplot(outDATA[[i]], col = c("darkblue", "darkred", "darkgreen"), outline =
FALSE,
#     main = nameLIST[[i]], cex.main = 0.8, names = FALSE, las = 2, log = "y", ylim =
YLIM)
#     CONF <- apply(outDATA[[i]], 2, function(x) quantile(x, c(0.025, 0.975), na.rm =
TRUE))
#     COLS <- 1:ncol(CONF)
#     segments(COLS - 0.2, CONF[1, ], COLS + 0.2, CONF[1, ], col = c("darkblue",
"darkred", "darkgreen"), lwd = 2)
#     segments(COLS - 0.2, CONF[2, ], COLS + 0.2, CONF[2, ], col = c("darkblue",
"darkred", "darkgreen"), lwd = 2)

#   }
#   par(mar = c(0.5, 0.5, 0.5, 0.5))
#   plot(1, 1, type = "n", axes = FALSE)
#   legend(x = 0.7, y = 1.4, legend = c("Monte-Carlo\nSimulation", "Permutation",
"Error\nPropagation"),
#   bty = "n", cex = 1.3, fill = c("darkblue", "darkred", "darkgreen"), y.intersp = 2)
#   }
#   outFRAME2 <- cbind(VALS = rownames(outFRAME), outFRAME)
#   write.table(outFRAME2, file = ifelse(dataout == "clip", "clipboard-64000", dataout),
row.names = FALSE, sep = "\t")
#   return(list(resList = outLIST, resDat = outFRAME, rawData = outDATA))

```

8.2.1.2 R-Skript zum Ermitteln der Referenzgene

```
#-----  
# geNorm.R  
#  
# Analysiert berechnete Cp und Effizienz zur Auswertung stabil  
# replizierter Gene (Referenzgene)  
# bewertet und vergleicht die Gene paarweise und ermittelt  
# Stabilitätsfaktor M  
#  
# Daten Vorverarbeitung:  
# Export der qpcr.R-Daten und Bearbeitung der Effizienzen als .csv  
# (comma separated values)  
#  
# Voraussetzungen:  
# R Version > 2.13 (http://cran.r-project.org/)  
# SLqPCR package (http://bioconductor.org/biocLite.R)  
#-----  
library(SLqPCR)  
  
pdf("kmmXX.pdf",width=8,height=5,pointsize=4,paper="special")  
  
par(mfrow=c(1,2))  
  
csvfile <- "kmmXX_geNorm_alle.csv"  
# csvfile <- "relWertekmm.csv"  
  
refdata <- read.csv2(csvfile, sep="," , dec="," )  
  
# data(vandesompele)  
# str(refdata)  
# laden der relativen Werte  
  
tissue <- as.factor(c(rep("Vgl", 12)))  
res.Vgl <- selectHKgenes(refdata[1:12,], method = "Vandesompele",  
geneSymbol = names(refdata), minNrHK = 2, trace = TRUE, na.rm = TRUE)  
  
# schrittweise definieren der HKG:  
# dabei werden die instabilsten Gene  
#- höchster M-Wert - nach und nach ausgeschlossen,  
# solange bis nur noch 2 minNrHK übrig sind  
  
ranks <- data.frame(c(1, 1:9), res.Vgl$ranking)  
names(ranks) <- c("rank", "Vgl")  
ranks  
  
#exa1 <- apply(refdata[tissue == "B.megaterium", ], 2, relQuantPCR, E = 2.0)
```

```
library(RColorBrewer)
mypalette <- brewer.pal(5, "Set1")
matplot(cbind(res.Vgl$meanM), type = "b",
ylab = "Average expression stability M",
xlab = "Number of remaining control genes", axes = FALSE, pch = 19,
col = mypalette,
ylim = c(0.2, 1.42), lty = 1, lwd = 2, main = "Gene stability measure")
axis(1, at = 1:9, labels = as.character(10:2))
axis(2, at = seq(0.2, 1.4, by = 0.2),
labels = as.character(seq(0.2, 1.4, by = 0.2)))
box()
abline(h = seq(0.2, 1.4, by = 0.2), lty = 2, lwd = 1, col = "grey")
legend("topright", legend = c("Vgl"), fill = mypalette)
# Graphik 1
```

```
mypalette <- brewer.pal(8, "YlGnBu")
barplot(cbind(res.Vgl$variation), beside = TRUE, col = mypalette,
space = c(0, 2), names.arg = c("Vgl"))
legend("topright", legend = c("V9/V10", "V8/V9", "V7/V8", "V6/V7",
"V5/6", "V4/5", "V3/4", "V2/3"), fill = mypalette, ncol = 2)
abline(h = seq(0.05, 0.25, by = 0.05), lty = 2, col = "grey")
abline(h = 0.15, lty = 1, col = "black")
# Graphik 2
```

```
# Normalisieren
data(SLqPCRdata)
SLqPCRdata
(relData <- apply(SLqPCRdata, 2, relQuantPCR, E = 2))
geneStabM(relData[, c(3, 4)])
(exprData <- normPCR(SLqPCRdata, c(3, 4)))

dev.off()

# dat2rest
# ratiocalc
```

8.2.2 Microarray

8.2.2.1 Auswertung des DNA-Microarrays Teil 1

```
#-----
# Bacillus megaterium GFP production via xylose unducible system
#
# Analysiert Daten des Microarray-Scanner (Agilent)
# Korrektur über Hintergrund möglich
```

```
#
# Daten Vorverarbeitung:
# definieren der targets aus den Slides in target-file
# (mit Sample Number, FileNames, Cy3, Cy5, Date -> targets.txt)
#
# Voraussetzungen:
# R Version > 2.13 (http://cran.r-project.org/)
# limma package (http://bioconductor.org/biocLite.R)
# Biobase package (http://bioconductor.org/biocLite.R)
# gplots package (http://cran.r-project.org/)
#-----

## load libraries
library("limma")
library("Biobase")
library("gplots")

## set working directory
targetfile.dir <- "/home/sheyber/Master/Array"
setwd(targetfile.dir)

## filename of the R file
filename <- "bmeg_induction.R"

## Name of the spot type-file
spotTypesFile <- "SpotTypes.txt"

## Name of the description File
targetFile <- "targets.txt"
reference <- "ref"
## set log(2) fold change cutoff
logfc_cutoff <- 0.8

## set path separator
pathSep <- "/"

#####
###INPUT###
#####
## read target information
targets <- readTargets("targets.txt")

## read-in with quality markers
RGraw <- read.maimages(targets, path=targetfile.dir, columns =
list(G = "gMedianSignal", Gb = "gBGMedianSignal",
R = "rMedianSignal",
Rb = "rBGMedianSignal"), annotation =
c("SystematicName", "ProbeName", "GeneName", "Description",
```

```
"ControlType", "Row", "Col",
"glIsSaturated", "rlIsSaturated", "glFeatNonUnifOL",
"rlFeatNonUnifOL", "glBGNonUnifOL", "rlBGNonUnifOL",
"glIsPosAndSignif", "rlIsPosAndSignif", "glIsWellAboveBG",
"rlIsWellAboveBG"), source = "agilent")

## microarray layout for spatial/quality plots
colnames(RGraw$genes)[which(colnames(RGraw$genes) == "Col")] <- "Column"
RGraw$genes$Block <- 1
RGraw$printer = getLayout(RGraw$genes)

## read spot types from file SpotTypes.txt
spottypes <- readSpotTypes(file = spotTypesFile, sep = "\t")
RGraw$genes$Status <- controlStatus(spottypes, RGraw)

## calculate MA values
MAraw <- MA.RG(RGraw)

## remove control spots from the RGList
controls <- which(RGraw$genes$Status %in% c("posControl", "negControl"))
RGraw.woc <- RGraw[-controls, ]
MAraw.woc <- MAraw[-controls, ]

# save raw data to object/load raw data from object
save(targets, RGraw, MAraw, spottypes, controls, RGraw.woc, MAraw.woc,
file = "rawdata.RData")
load("rawdata.RData")

#####
###BACKGROUND-CORRECTION AND NORMALISATION###
#####

## background correction ##
RGraw.bgc <- backgroundCorrect(RGraw.woc, method = "normexp", offset = 50)

## normalize within arrays ##
## normalization method can be specified by "method" parameter:
## "none", "median", "loess", "printtiploess", "composite",
## "control" and "robustspline"
## default for loess is span = 0.3, lo(w)ess
## normalisation is appropriate for an even distribution of up and down
## regulated genes (-> MA-Plot) and for maximal
## 20% to 30% differentiel expressed genes as a concervative estimation
## normalizeWithinArrays can perform background correction with default:
## bc.method="subtract" and offset=0
MAnorm.wa <- normalizeWithinArrays(RGraw.bgc, method = "loess")

## if processed signal values were read-in
```

```
#MAnorm.wa <- normalizeWithinArrays(RGraw.bgc, method = "none")

## normalize between arrays ##
## "quantile" normalization (Yang and Thorne (2003)) ensures that the
## intensities have the same empirical distribution across arrays
## and across channels
## "scale" normalization (Yang et al (2001, 2002), Smyth and Speed (2003))
## scales the log-ratios to have the same
## median-absolute-deviation (MAD) across arrays
MAnorm.waba <- normalizeBetweenArrays(MAnorm.wa, method = "quantile")

## save normalized data to object/load normalized data from object ##
save(RGraw.bgc, MAnorm.wa, MAnorm.waba, file = "normdata.RData")
load("normdata.RData")

#####
###DIFF. EXPRESSED GENES AND DATAEXPORT###
#####

## average replicate spots ##
MAnorm.avg <- avereps(MAnorm.waba, ID=MAnorm.waba$genes$SystematicName)

## build design matrix using 'ref' as reference
design <- modelMatrix(targets, ref = reference)

## lmFit + eBayes
# -> linear model for the calculation/estimation of standard errors and
## other parameter for the statistical tests
# -> computation of statistical tests
LMres <- lmFit(MAnorm.avg, weights = NULL, design = design)
LMres <- eBayes(LMres)

# save model+design data to object/load model+design data from object
save(LMres, design, file="model.RData")
load("model.RData")

## Dataexport ##
restable <- list()
restable2 <- list()
restable21 <- list()
restable3 <- list()
restable31 <- list()
for (i in 1:ncol(LMres$coefficients))
{
  number <- length(LMres$coefficients[, i])
  name <- colnames(LMres$coefficients)[i]
  #first preliminary result table
```

```

restable[[name]] <- topTable(LMres, number = number, coef = i)

## order without cutoff and export
# order by logFC value
restable2[[name]] <- restable[[name]][order(restable[[name]]$logFC, decreasing =
TRUE), ]
restable2[[name]]["FC"] <- 2^abs(restable2[[name]]["logFC"])
tmp <- restable2[[name]]["logFC"] > 0
tmp[tmp == c(FALSE)] <- -1
restable2[[name]]["FC"] <- restable2[[name]]["FC"]*tmp
restable21[[name]] <- restable2[[name]][c("SystematicName", "GeneName",
"Description", "ProbeName", "Row", "Column", "logFC", "FC", "P.Value", "adj.P.Val")]
write.table(restable21[[name]], file = paste(filename, "_Genelist_all-", name, ".tab",
sep=""),
sep="\t", dec=".", row.names=FALSE)
## set logFC cutoff and export

restable3[[name]] <- subset(restable2[[name]], abs(restable2[[name]]$logFC)
>= logfc_cutoff)
restable3[[name]]["FC"] <- 2^abs(restable3[[name]]["logFC"])
tmp <- restable3[[name]]["logFC"] > 0
tmp[tmp == c(FALSE)] <- -1
restable3[[name]]["FC"] <- restable3[[name]]["FC"]*tmp
restable31[[name]] <- restable3[[name]][c("SystematicName", "GeneName",
"Description", "ProbeName", "Row", "Column", "logFC", "FC", "P.Value", "adj.P.Val")]
write.table(restable31[[name]], file = paste(filename, "_Genelist_logFC",
logfc_cutoff, "-", name, ".tab", sep=""), sep="\t", dec=".", row.names=FALSE)
}

# save result tables to object/load result tables from object
save(restable, restable2, restable3, file = "restables.Rdata")

```

8.2.2.2 Auswertung des DNA-Microarrays Teil 2

```

#-----
# Bacillus megaterium GFP production via xylose unducible system
#
# Erstellt aus den analysierten Daten des Microarrays
# erstellt Heatmap der (hier) 50 am höchsten differentiell
# exprimierten Gene
#
# Daten Vorverarbeitung:
# Erstellung der induzierten Gene als .csv
#
# Voraussetzungen:

```

```
# R Version > 2.13 (http://cran.r-project.org/)
# gplots package (http://cran.r-project.org/)
#-----

library(gplots)

pdf("bmeg_cluster_heatmap.pdf", height=15, width=15)

# Read microarray time series data
csvfile <- "bmeg_induction_all.csv"
bm <- read.csv(csvfile, sep="\t", dec=".")

# generate unique IDs Vector (use orf ID if gene name not available)
GeneNameVector <- as.vector(bm$GeneName)
SysNameVector <- as.vector(bm$SystematicName)
GeneNameVector[which(GeneNameVector == "")]
<- SysNameVector[which(GeneNameVector=="")]
bm <- cbind(bm, data.frame(IDs = GeneNameVector))

top_no <- 50 # top number of differentially expressed genes

# extract data for clustering
cluster.data <- data.matrix(data.frame(bm$logFC30, bm$logFC60, bm$logFC120)
[1:top_no,])
row_labels <- as.vector(bm$IDs[1:top_no])
col_labels <- c("0.5h", "1h", "2h")

# draw heatmap
heatmap.2(cluster.data,
  dendrogram = "row",
  Colv = FALSE,
  col = redgreen(256),
  # col = bluered(256),
  scale = "none",
  # key = TRUE,
  key = FALSE,
  density.info = "none",
  trace = "none",
  margins = c(8,20),
  cexRow = 1.5,
  cexCol = 2.5,
  lhei = c(0.05,0.95),
  labRow = row_labels,
  labCol = col_labels)
dev.off()

#-----
# Bacillus megaterium GFP production via xylose unducible system
```



```
#
# Erstellt aus den analysierten Daten des Microarrays
# VennDiagramm mit Unterscheidung der zu einem Zeitpunkt induzierten Gene
# zusätzlich: barplot aller induzierten Gene
#
# Daten Vorverarbeitung:
# Erstellung der induzierten Gene als .csv
# Definition cut-off-Value (hier: 1.59)
#
# Voraussetzungen:
# R Version > 2.13 (http://cran.r-project.org/)
# VennDiagram package (http://cran.r-project.org/)
#-----

library(VennDiagram)

# Read microarray time series data
csvfile <- "bmeg_induction_all.csv"
bm <- read.csv(csvfile, sep="\t", dec=".")
T05_amount <- as.vector(unlist(subset(bm, abs(logFC30) > 1.59,
select="SystematicName"))))
T1_amount <- as.vector(unlist(subset(bm, abs(logFC60) > 1.59,
select="SystematicName"))))
T2_amount <- as.vector(unlist(subset(bm, abs(logFC120) > 1.59,
select="SystematicName"))))

venn.diagram(
x = list(
T0.5h = T05_amount,
T1h = T1_amount,
T2h = T2_amount
),
filename = "venn_diagram.tiff",
col = "black",
lty = "dotted",
lwd = 4,
fill = c("cornflowerblue", "green", "yellow"),
alpha = 0.50,
label.col = c("black", "black", "black", "black", "black", "black", "black"),
cex = 2.5,
fontfamily = "sansserif",
fontface = "bold",
# cat.col = c("darkblue", "darkgreen", "orange"),
cat.col = c("black", "black", "black"),
cat.cex = 2.5,
cat.fontfamily = "sansserif"
)
```

```
pdf("diffgenes.pdf", height=5, width=7)
barplot(c(length(T05_amount),length(T1_amount),length(T2_amount)),
names.arg = c("T05h", "T1h", "T2h"), ylim = c(0,300),
col="cornflowerblue", main="Differentially expressed genes")
dev.off()
```

8.2.3 Auswertung der differentiell exprimierten Gene

```
library(gplots)

pdf("bmeg_cluster_heatmap.pdf", height=15, width=15)

# Read microarray time series data
csvfile <- "bmeg_induction_all.csv"
bm <- read.csv(csvfile, sep="\t", dec=".")

# generate unique IDs Vector (use orf ID if gene name not available)
GeneNameVector <- as.vector(bm$GeneName)
SysNameVector <- as.vector(bm$SystematicName)
GeneNameVector[which(GeneNameVector == "")] <-
SysNameVector[which(GeneNameVector=="")]
bm <- cbind(bm, data.frame(IDs = GeneNameVector))

top_no <- 50 # top number of differentially expressed genes

# extract data for clustering
cluster.data <- data.matrix(data.frame(bm$logFC30, bm$logFC60, bm$logFC120)
[1:top_no,])
row_labels <- as.vector(bm$IDs[1:top_no])
col_labels <- c("0.5h", "1h", "2h")

# draw heatmap
heatmap.2(cluster.data,
          dendrogram = "row",
          Colv = FALSE,
          col = redgreen(256),
          # col = bluered(256),
          scale = "none",
          # key = TRUE,
          key = FALSE,
          density.info = "none",
          trace = "none",
          margins = c(8,20),
          cexRow = 1.5,
          cexCol = 2.5,
          lhei = c(0.05,0.95),
```

```
labRow = row_labels,  
labCol = col_labels)
```

```
dev.off()
```

8.3 Auszug der differentiell exprimierten Gene

Die folgende Zusammenstellung beinhaltet basierend auf den Auswertungen des DNA-Microarrays einen Auszug der differentiell exprimierten Gene. Jeder Zeitpunkt der Probenahme (30, 60 und 120 min nach Induktion) ist als \log_2 -Expressionsverhältnis aufgeführt. Als Grenzwert wurde ein absolutes Expressionsverhältnis von mindestens 3-fach zu allen drei Zeitpunkten definiert ($|\log_2(FC)| \geq 1$). Hier soll ein absolutes Expressionsverhältnis bis einschließlich 6-fach gezeigt werden, als Ausnahme wurde der Transporter zugefügt.

Systematic Name	Gene Name	Description	FC30	FC60	FC120	FCScore
BMD_0684	argF	ornithine carbamoyltransferase	14,98	10,20	5,95	31,13
BMD_4657	maf	septum formation protein Maf	5,05	13,42	7,63	26,09
BMD_2932		hypothetical protein	3,19	10,27	12,46	25,92
BMD_1858	xylA	xylose isomerase	5,20	15,97	4,18	25,35
BMD_2605	cblX	sirohydrochlorin cobaltochelataase	-5,13	-13,45	-5,96	-24,53
BMD_4656	radC	DNA repair protein RadC	5,08	13,43	5,03	23,54
BMD_1450	cspD	cold shock protein	2,91	9,09	11,49	23,48
BMD_4063	gabT	4-aminobutyrate transaminase	-5,06	-11,60	-6,55	-23,21
BMD_2208		general stress protein 17M	4,36	10,15	8,63	23,15
GFP_62	GFP_62	Unknown	2,24	6,64	14,14	23,02
BMD_1857	xylR	xylose repressor	3,40	8,19	11,40	22,99
BMD_2897		conserved hypothetical protein	3,09	4,82	14,34	22,25
GFP_171	GFP_171	Unknown	2,03	4,26	14,41	20,71
BMD_1557		glucose starvation-inducible protein B	2,51	4,73	12,76	19,99
BMD_4776	argG	argininosuccinate synthase	7,28	9,27	2,97	19,52
BMD_2603	cblC	precorrin-8X methylmutase CblC	-4,42	-10,19	-4,73	-19,33
BMD_4713	lysC	aspartate kinase	1,31	8,75	8,44	18,50
BMD_3013		hypothetical protein	1,58	3,63	13,23	18,45
BMD_4775	argH	argininosuccinate lyase	7,95	7,81	2,63	18,39
GFP_203	GFP_203	Unknown	1,87	4,68	11,83	18,38
BMD_3145		hypothetical protein	3,90	9,21	4,51	17,61
BMD_0683	carB	carbamoyl-phosphate synthase, large subunit	7,85	6,88	2,65	17,38
BMD_0904		hypothetical protein	2,25	5,62	9,45	17,32
BMD_0681	argD	acetylornithine aminotransferase	7,72	5,90	2,89	16,52
BMD_2606	cblH	precorrin 3 methylase	-4,03	-8,23	-4,15	-16,40
BMD_1682	cspA	cold shock protein	2,85	5,82	7,22	15,89
GFP_229	GFP_229	Unknown	1,69	3,68	10,37	15,74

Populationsheterogenität während der heterologen Proteinproduktion in *Bacillus megaterium*

BMD_2600	cbiL	precorrin-2 C20-methyltransferase	-3,77	-8,42	-3,48	-15,67
BMD_0680	argB	acetylglutamate kinase	7,38	6,09	2,12	15,60
BMD_0148	rpsH	30S ribosomal protein S8	-1,96	-4,00	-9,12	-15,08
BMD_3693	bioF	8-amino-7-oxononanoate synthase	3,18	8,00	3,68	14,86
BMD_2599	cbiF	precorrin-4 C11-methyltransferase	-3,78	-7,50	-3,52	-14,81
BMD_2910		conserved hypothetical protein	-2,25	-6,63	-5,80	-14,68
BMD_0958	treC	alpha,alpha-phosphotrehalase	-3,68	-7,05	-3,92	-14,65
BMD_2601	cbiET	CbiET protein	-3,70	-7,59	-3,18	-14,47
GFP_145	GFP_145	Unknown	1,59	2,63	10,20	14,42
BMD_0135	rplD	50S ribosomal protein L4	-1,62	-2,98	-9,67	-14,27
BMD_0682	carA	carbamoyl-phosphate synthase, small subunit	6,92	4,03	3,29	14,25
BMD_2285	nprM	Bacillolysin precursor (neutral protease)	2,27	6,52	5,40	14,18
BMD_4636	nadA	quinolinate synthetase complex, A subunit	1,18	7,92	4,91	14,01
BMD_2602	cbiD	cobalamin biosynthesis protein CbiD	-3,54	-7,28	-2,99	-13,80
BMD_3867	nucB	sporulation-specific extracellular nuclease	3,36	6,00	4,44	13,80
BMD_0957	scrA	PTS system trehalose-specific EIIBC component	-3,52	-6,43	-3,55	-13,50
BMD_1178		conserved hypothetical protein	3,32	6,38	3,68	13,38
BMD_3090		conserved hypothetical protein	3,08	5,07	5,14	13,29
BMD_4638	nadB	L-aspartate oxidase	1,40	6,03	5,75	13,18
BMD_4416	artM	arginine ABC transporter, ATP-binding protein ArtM	4,56	5,65	2,87	13,08
SPYocH	SPYocH	Unknown	1,35	1,59	10,14	13,08
BMD_5137	atpH	ATP synthase F1, delta subunit	-2,66	-4,66	-5,75	-13,07
BMD_4049		Transporter (Major facilitator Superfamily)	-3,14	-5,72	-4,15	-13,01
BMD_4745	speH	S-adenosylmethionine decarboxylase	2,49	5,32	5,19	13,00
BMD_1479		conserved hypothetical protein	3,26	7,45	2,27	12,97
BMD_3217	yclP	putative ferrichrome ABC transporter, ATP-binding protein	-3,14	-5,50	-4,32	-12,96
BMD_0199		glyoxalase family protein	2,22	4,38	6,34	12,94
BMD_2604	cbiJ	precorrin-6x reductase	-3,35	-7,00	-2,58	-12,93
BMD_4744		conserved hypothetical protein	2,14	5,32	5,42	12,88
BMD_0641	fabG	3-oxoacyl-(acyl-carrier-protein) reductase	1,83	3,91	7,12	12,86
BMD_0150	rplR	50S ribosomal protein L18	-1,84	-3,59	-7,32	-12,75
BMD_0149	rplF	50S ribosomal protein L6	-1,84	-3,23	-7,66	-12,73
GFP_34	GFP_34	Unknown	1,44	2,41	8,78	12,63
BMD_1477	cls	cardiolipin synthetase	2,88	7,61	2,10	12,59
BMD_1321		conserved hypothetical protein	-2,02	-4,71	-5,84	-12,57
BMD_4552	grpE	co-chaperone GrpE	-1,98	-4,63	-5,89	-12,49
BMD_2791	cspC	cold shock protein	2,55	4,73	5,07	12,35
BMD_0130	rpsG	30S ribosomal protein S7	-1,36	-2,73	-8,24	-12,33
BMD_0186		hypothetical protein	2,15	4,91	5,20	12,25
BMD_0147	rpsN	30S ribosomal protein S14	-1,87	-3,43	-6,94	-12,25
BMD_0138	rpsS	30S ribosomal protein S19	-1,83	-3,53	-6,83	-12,19
BMD_0146	rplE	50S ribosomal protein L5	-1,91	-3,56	-6,68	-12,15

Populationsheterogenität während der heterologen Proteinproduktion in *Bacillus megaterium*

BMD_0136	rplW	50S ribosomal protein L23	-1,64	-2,68	-7,76	-12,09
BMD_0292	gatC	glutamyl-tRNA(Gln) and/or aspartyl-tRNA(Asn)	-2,04	-5,12	-4,89	-12,05
BMD_4062		amidotransferase, C subunit				
BMD_4637	nadC	Sigma-54 interaction domain	-3,39	-5,43	-3,11	-11,93
BMD_0129	rpsL	nicotinate-nucleotide diphosphorylase (carboxylating)	1,21	7,21	3,50	11,92
BMD_4061		30S ribosomal protein S12	-1,44	-2,88	-7,60	-11,92
		succinate-semialdehyde dehydrogenase [NADP+]	-3,09	-4,95	-3,85	-11,90
GFP_100	GFP_100	Unknown	1,44	2,19	8,26	11,88
BMD_0638		conserved hypothetical protein	1,83	4,42	5,62	11,87
BMD_1249	clpE	ATP-dependent Clp protease, ATP- binding subunit ClpE	2,73	4,85	4,15	11,73
BMD_5138	atpF	ATP synthase F0, B subunit	-2,47	-4,03	-5,16	-11,66
BMD_3219	yclN	putative ferrichrome ABC transporter, permease protein	-2,93	-6,41	-2,27	-11,62
BMD_3372		hypothetical protein	1,74	3,60	6,26	11,60
BMD_5086		sigma 54 modulation protein / S30EA ribosomal protein	1,74	5,53	4,33	11,60
BMD_2597	cbiA	cobyrinic acid a,c-diamide synthase	-2,97	-5,83	-2,79	-11,59
BMD_3748		conserved hypothetical protein	1,71	3,24	6,58	11,53
BMD_0142	rpmC	50S ribosomal protein L29	-1,84	-3,26	-6,34	-11,44
BMD_4309	trpD	anthranilate phosphoribosyltransferase	1,88	6,63	2,86	11,37
BMD_4307	trpF	N-(5'-phosphoribosyl) anthranilate isomerase	1,90	6,34	3,14	11,37
BMD_1703	nikB	nickel import ABC transporter, permease subunit NikB	-2,67	-4,28	-4,38	-11,33
BMD_4857	dps	DNA-protecting protein	1,98	3,42	5,92	11,33
BMD_0133	rpsJ	30S ribosomal protein S10	-1,58	-2,71	-7,01	-11,30
BMD_1297		conserved hypothetical protein	2,37	5,06	3,87	11,29
BMD_2218		hypothetical protein	1,92	4,06	5,31	11,28
BMD_1700		glutamate dehydrogenase	-2,12	-3,99	-5,15	-11,26
BMD_0894		conserved hypothetical protein	3,40	4,25	3,56	11,21
BMD_0088	hpt	hypoxanthine phosphoribosyltransferase	-2,02	-4,26	-4,93	-11,20
BMD_1859	xylB	xylulokinase	3,22	5,89	1,94	11,05
BMD_4872	nrdE	ribonucleoside-diphosphate reductase, alpha subunit	-2,56	-4,93	-3,52	-11,02
BMD_0790	alsS	acetolactate synthase, catabolic	2,73	4,10	4,13	10,96
BMD_0701	appA	oligopeptide ABC transporter, oligopeptide-binding protein AppA	2,08	4,77	4,04	10,89
BMD_0971		hypothetical protein	1,91	4,54	4,37	10,82
BMD_3075		Diacylglycerol kinase catalytic domain protein	2,23	3,93	4,61	10,77
BMD_2614		IDEAL domain protein	1,47	3,90	5,38	10,74
BMD_4760	pyk	pyruvate kinase	-1,71	-3,62	-5,39	-10,73
BMD_1376		conserved hypothetical protein	-1,95	-4,77	-3,99	-10,71
BMD_2595	cbiY	nitroreductase family protein	-2,90	-5,40	-2,40	-10,70
BMD_0140	rpsC	30S ribosomal protein S3	-1,78	-3,36	-5,53	-10,67
BMD_3353	fabI	enoyl-[acyl-carrier protein] reductase I	-1,92	-4,05	-4,49	-10,45

BMD_0679	argJ	arginine biosynthesis bifunctional protein ArgJ	5,74	3,21	1,50	10,45
BMD_4960	thrB	homoserine kinase	1,17	6,97	2,29	10,43
BMD_1890		conserved hypothetical protein	2,12	4,40	3,83	10,36
BMD_4132	pnp	polynucleotide phosphorylase (PNPase)	-1,68	-3,53	-5,12	-10,33
BMD_5053	hisF	imidazole glycerol phosphate synthase, cyclase subunit	-1,02	5,50	3,79	8,26
BMD_3142		hypothetical protein	2,37	5,09	2,82	10,28
BMD_5140	atpB	ATP synthase F0, A subunit	-2,13	-3,25	-4,90	-10,28
BMD_5135	atpG	ATP synthase F1, gamma subunit	-2,62	-4,50	-3,10	-10,21
BMD_0143	rpsQ	30S ribosomal protein S17	-1,86	-3,36	-4,99	-10,20
BMD_1049		hypothetical protein	1,74	3,64	4,81	10,19
BMD_4306	trpB	tryptophan synthase, beta subunit	1,83	5,82	2,53	10,18
BMD_5054	hisA	phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase	-1,02	6,53	2,63	8,14
BMD_0089	ftsH	cell division protease FtsH	-1,77	-3,45	-4,95	-10,17
BMD_5000	yfiY	putative ferrichrome import ABC transporter, ferrichrome-binding protein	-3,06	-4,15	-2,92	-10,13
BMD_0141	rplP	50S ribosomal protein L16	-1,78	-3,14	-5,16	-10,08
BMD_1062		Proton/sodium-glutamate symport protein	-2,79	-4,12	-3,16	-10,08
BMD_0131	fusA	translation elongation factor G (EF-G)	-1,41	-2,81	-5,82	-10,05
BMD_3039	lytF	cell wall endopeptidase	-2,16	-4,35	-3,54	-10,04
BMD_0869		acyl-CoA dehydrogenase	1,58	3,12	5,33	10,03
BMD_1781		glucose starvation-inducible protein B (General stress protein B)	1,77	2,68	5,57	10,02
BMD_0945		putative 4-aminobutyrate transaminase	4,40	3,53	2,05	9,98
BMD_1379	asnS	asparaginyl-tRNA synthetase	-1,71	-3,32	-4,93	-9,95
BMD_1326	pdhA	pyruvate dehydrogenase E1 component, alpha subunit	-1,96	-3,68	-4,32	-9,95
BMD_1181		conserved hypothetical protein	-1,85	-4,51	-3,57	-9,93
BMD_5136	atpA	ATP synthase F1, alpha subunit	-2,41	-3,83	-3,67	-9,91
BMD_3216	yclQ	putative ferrichrome ABC transporter, ferrichrome-binding protein	-2,92	-4,01	-2,96	-9,90
BMD_1478		protein of unknown function (DUF1206)	2,60	5,51	1,78	9,89
BMD_1327	pdhB	pyruvate dehydrogenase E1 component, beta subunit	-1,83	-3,27	-4,74	-9,85
BMD_4613	hisS	histidyl-tRNA synthetase	-1,78	-3,39	-4,67	-9,84
BMD_0155	adk	adenylate kinase	-1,69	-3,21	-4,90	-9,80
BMD_0137	rplB	50S ribosomal protein L2	-1,53	-2,36	-5,84	-9,72
BMD_2273		conserved hypothetical protein	2,77	4,64	2,31	9,72
BMD_1702	nikA	nickel import ABC transporter, nickel-binding protein NikA	-2,54	-3,75	-3,39	-9,68
BMD_2648		conserved hypothetical protein	1,81	3,39	4,43	9,63
BMD_0156	map	methionine aminopeptidase, type I	-1,56	-2,82	-5,23	-9,61
BMD_0144	rplN	50S ribosomal protein L14	-1,76	-3,11	-4,72	-9,59

BMD_0293	gatA	glutamyl-tRNA(Gln) and/or aspartyl-tRNA(Asn) amidotransferase, A subunit	-1,75	-3,33	-4,51	-9,59
BMD_3305	gvpP	gas vesicle protein GvpP	1,68	2,26	5,64	9,57
BMD_1121		polysaccharide biosynthesis protein	-1,77	-3,79	-3,87	-9,43
BMD_4305	trpA	tryptophan synthase, alpha subunit	1,71	4,98	2,73	9,42
BMD_0678	argC	N-acetyl-gamma-glutamyl-phosphate reductase	4,87	3,04	1,50	9,41
BMD_2074		two-component response regulator	1,77	2,69	4,95	9,41
BMD_5036	tpiA	triosephosphate isomerase	-2,17	-3,67	-3,55	-9,39
BMD_0478	yocH	putative exported cell wall-binding protein	-2,00	-3,15	-4,22	-9,37
BMD_5197	gnd	6-phosphogluconate dehydrogenase, decarboxylating	-1,74	-2,66	-4,96	-9,36
BMD_4053		L-2,4-diaminobutyrate decarboxylase	-2,70	-3,92	-2,73	-9,36
BMD_2499	ppaC	manganese-dependent inorganic pyrophosphatase	-1,70	-2,38	-5,28	-9,36
BMD_4308	trpC	indole-3-glycerol-phosphate synthase	1,60	5,03	2,72	9,34
BMD_0157	infA	translation initiation factor IF-1	-1,55	-2,62	-5,17	-9,33
BMD_2513		HAD-superfamily hydrolase, subfamily IA, variant 3	-1,95	-3,24	-4,13	-9,32
BMD_1761		conserved hypothetical protein	2,13	3,58	3,58	9,29
BMD_5199	ald	alanine dehydrogenase	-2,35	-3,90	-3,03	-9,28
BMD_4261	ftsA	cell division protein FtsA	-1,76	-3,11	-4,37	-9,23
BMD_1058		hypothetical protein	1,65	3,17	4,40	9,23
BMD_4048		siderophore biosynthesis protein	-2,60	-3,80	-2,78	-9,18
BMD_4979	sufD	FeS assembly protein SufD	-1,85	-3,00	-4,31	-9,15
BMD_4980	sufC	FeS assembly ATPase SufC	-1,78	-3,10	-4,25	-9,13
BMD_2860		conserved hypothetical protein	2,22	3,24	3,66	9,12
BMD_4192	sucC	succinate-CoA ligase, beta subunit	-2,29	-2,60	-4,21	-9,10
BMD_2466		Allophanate hydrolase subunit 1	-1,73	-4,94	-2,41	-9,08
BMD_4052		siderophore biosynthesis protein	-2,75	-3,61	-2,70	-9,06
BMD_2594	btuR	cob(I)alamin adenosyltransferase	-2,45	-4,29	-2,31	-9,06
BMD_5165	rpoE	DNA-directed RNA polymerase, delta subunit	-1,49	-3,14	-4,41	-9,04
BMD_3064		transition state regulator, domain protein	-2,13	-4,05	-2,85	-9,03
BMD_1538		Ferritin-like domain protein	2,04	4,09	2,88	9,00
BMD_0152	rpmD	50S ribosomal protein L30	-1,62	-2,72	-4,65	-8,99
BMD_4086	glnA	glutamine synthetase, type I	-1,87	-3,05	-4,05	-8,96
BMD_3698		sugar-specific transcriptional regulator TrmB family	1,59	2,82	4,55	8,96
BMD_0145	rplX	50S ribosomal protein L24	-1,85	-3,41	-3,69	-8,95
BMD_3868	nin	DNA-entry nuclease inhibitor (competence protein J)	2,62	4,09	2,22	8,93
BMD_0418		conserved hypothetical protein	2,17	4,58	2,18	8,93
BMD_3167		conserved hypothetical protein	2,48	3,95	2,50	8,93
BMD_4032		FAD-dependent dehydrogenase	3,04	4,18	1,66	8,88
BMD_1328	pdhC	pyruvate dehydrogenase complex E2 component, dihydrolipoamide acetyltransferase	-1,86	-3,08	-3,92	-8,87

BMD_4417	artQ	arginine ABC transporter, permease protein ArtQ	3,50	3,79	1,56	8,85
BMD_4310	trpE	anthranilate synthase component I	1,59	4,36	2,88	8,84
BMD_4351	serA	D-3-phosphoglycerate dehydrogenase	1,92	3,93	2,96	8,81
BMD_1373	panD	aspartate 1-decarboxylase	1,18	5,69	1,94	8,81
BMD_4948		putative membrane protein	1,55	2,42	4,83	8,80
BMD_0139	rplV	50S ribosomal protein L22	-1,84	-3,42	-3,52	-8,78
BMD_1268		guanine-hypoxanthine permease	2,46	3,70	2,61	8,78
BMD_2099		hypothetical protein	2,71	3,84	2,21	8,76
BMD_3695	bioC	biotin biosynthesis protein BioC	2,80	4,80	-1,12	6,48
BMD_2747		glycerate kinase family protein	1,31	2,16	5,24	8,71
BMD_3121	cysI	sulfite reductase (NADPH) hemoprotein, beta-component	-1,95	-3,11	-3,62	-8,69
BMD_4608	pstA	phosphate ABC transporter, permease protein PstA	1,42	2,18	5,09	8,69
BMD_0521		conserved hypothetical protein	1,32	3,07	4,28	8,68
BMD_1796		glucose 1-dehydrogenase	2,07	3,67	2,93	8,67
BMD_4367	ppiB	peptidyl-prolyl cis-trans isomerase, cyclophilin-type	-1,67	-2,96	-4,01	-8,65
BMD_2596	cysG	uroporphyrin-III C-methyltransferase	-2,32	-3,83	-2,49	-8,64
BMD_3649		hydrolase, alpha/beta fold family	2,33	4,07	2,22	8,62
BMD_4073	tkt	transketolase	-1,71	-2,62	-4,27	-8,60
BMD_1120		glycosyl transferase, group 1	-1,66	-2,76	-4,17	-8,59
BMD_3907	pepQ	Xaa-Pro dipeptidase	-1,76	-3,19	-3,63	-8,59
BMD_0154	secY	preprotein translocase, SecY subunit	-1,58	-3,02	-3,98	-8,58
BMD_0395		amidohydrolase	2,15	3,47	2,94	8,57
BMD_4743	nrdR	transcriptional regulator NrdR	1,94	4,17	2,44	8,55
BMD_3006		ThiJ/PfpI family protein	1,99	3,76	2,81	8,55
BMD_2156		probable transcription regulator sinR	-2,08	-4,63	-1,82	-8,53
BMD_5055	hisH	imidazole glycerol phosphate synthase, glutamine amidotransferase subunit	-1,21	5,22	2,09	6,10
BMD_4096		amino acid transporter	1,64	3,80	3,07	8,51
BMD_1118		glycosyl transferase, family 2	-1,62	-2,97	-3,91	-8,50
BMD_4186	hslV	ATP-dependent protease HslV	-1,67	-2,56	-4,27	-8,50
BMD_4425	bkdAA	2-oxoisovalerate dehydrogenase E1 component alpha subunit	-1,79	-3,08	-3,60	-8,47
BMD_0151	rpsE	30S ribosomal protein S5	-1,67	-2,86	-3,91	-8,45
BMD_2835		conserved hypothetical protein	1,69	3,21	3,55	8,44
BMD_0592	hit	protein Hit	-1,88	-3,01	-3,55	-8,44
BMD_2543	sspN	small acid soluble spore protein N	1,35	3,72	3,36	8,43
BMD_2607	cblW	cobalamin biosynthesis protein	-2,70	-4,47	-1,24	-8,41
BMD_3122		sulfite reductase [NADPH] flavoprotein, alpha-component	-1,97	-3,26	-3,17	-8,40
BMD_1372	panC	pantoate--beta-alanine ligase	1,10	6,20	-1,09	6,21
BMD_2464	accB	acetyl-CoA carboxylase, biotin carboxyl carrier protein	-1,73	-3,80	-2,84	-8,37
BMD_4961	thrC	threonine synthase	1,00	5,44	1,92	8,36

BMD_3218	yclO	putative ferrichrome ABC transporter, permease protein	-2,37	-4,08	-1,91	-8,36
BMD_4385	fur	ferric uptake regulation protein	2,54	4,14	1,67	8,35
BMD_0893		hypothetical protein	2,82	3,06	2,47	8,35
BMD_2228		RNA polymerase sigma factor, sigma-70 family	-1,69	-3,79	-2,87	-8,35
BMD_0111	gltX	glutamyl-tRNA synthetase	-1,60	-2,99	-3,72	-8,31
BMD_0577		conserved hypothetical protein	-2,32	-2,43	-3,52	-8,28
BMD_1342		putative membrane protein	-1,83	-3,72	-2,71	-8,26
BMD_4092	hfq	RNA chaperone Hfq	-1,77	-2,70	-3,79	-8,26
BMD_5191	gerQ	spore coat protein GerQ	1,45	2,25	4,55	8,25
BMD_4466		ribonucleoside-diphosphate reductase, adenosylcobalamin-dependent, frameshift	-1,67	-3,26	-3,33	-8,25
BMD_4241	pyrAA	carbamoyl-phosphate synthase, small subunit	-1,84	-3,23	-3,16	-8,23
BMD_0487	sspC	small acid soluble spore protein C	1,67	3,80	2,74	8,22
BMD_2100		sodium-dependent transporter	1,86	3,60	2,74	8,20
BMD_5229	amyL	alpha-amylase	2,04	3,34	2,82	8,20
BMD_1340	bipA	GTPase	-1,39	-2,56	-4,25	-8,20
BMD_4010	deoD	purine nucleoside phosphorylase	-1,40	-2,13	-4,66	-8,19
BMD_1253		oxidoreductase, short chain dehydrogenase/reductase family protein	-1,66	-3,29	-3,23	-8,18
BMD_4187	xerC	tyrosine recombinase XerC	-1,72	-2,76	-3,66	-8,15
BMD_4814		protein of unknown function (DUF1444)	-1,88	-3,41	-2,85	-8,14
BMD_5035	gpmI	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	-2,01	-3,37	-2,75	-8,13
BMD_5056	hisB	imidazoleglycerol-phosphate dehydratase	-1,21	5,13	1,78	5,70
BMD_1343		conserved hypothetical protein	-2,02	-3,30	-2,78	-8,10
BMD_2988	ureE	urease accessory protein UreE	-1,94	-3,23	-2,91	-8,07
BMD_3021	spo0M	sporulation-control protein Spo0M	-1,63	-2,85	-3,58	-8,07
BMD_4813		tRNA binding domain protein	-1,66	-2,99	-3,39	-8,04
BMD_1234	mtnW	2,3-diketo-5-methylthiopentyl-1-phosphate enolase	2,36	3,92	1,75	8,03
BMD_2677		Putative marR-family transcriptional regulator	-1,77	-3,74	-2,52	-8,03
BMD_5037	pgk	phosphoglycerate kinase	-1,82	-2,91	-3,27	-8,00
BMD_2282		conserved hypothetical protein	1,90	3,38	2,71	7,99
BMD_4528	sigA	RNA polymerase sigma factor	-1,76	-2,43	-3,80	-7,98
BMD_2284		hypothetical protein	-1,64	-3,54	-2,80	-7,98
BMD_1404	cspA	cold shock protein	-1,52	-1,91	-4,55	-7,98
BMD_3694	bioH	biotin biosynthesis protein BioH	2,54	4,11	1,32	7,98
BMD_1776		transglycosylase associated family protein	2,01	3,39	2,55	7,95
BMD_0158	rpsM	30S ribosomal protein S13	-1,46	-2,31	-4,18	-7,95
BMD_3892	ggt	gamma-glutamyltransferase	1,50	2,66	3,77	7,93
BMD_1463	ykfD	putative oligopeptide ABC transporter, ATP-binding protein	1,45	1,90	4,56	7,92
BMD_2119		helix-turn-helix domain protein	1,53	2,36	4,02	7,91

Populationsheterogenität während der heterologen Proteinproduktion in *Bacillus megaterium*

BMD_0260	groES	10 kDa chaperonin	-1,78	-2,16	-3,97	-7,91
BMD_0686		conserved hypothetical protein	1,65	2,63	3,61	7,89
BMD_3062		N-acetylmuramoyl-L-alanine amidase cwIB (Cell wall hydrolase) (Autolysin)	-1,76	-3,12	-3,00	-7,88
BMD_2231	ohrB	organic hydroperoxide resistance protein	2,26	2,65	2,97	7,88
BMD_4365		conserved hypothetical protein	2,30	3,28	2,31	7,88
BMD_0789	alsD	alpha-acetolactate decarboxylase	2,59	3,15	2,14	7,87
BMD_1364		pyrophosphatase YpjD	-1,74	-3,30	-2,82	-7,86
BMD_2546	acnA	aconitate hydratase 1	-1,57	-2,11	-4,17	-7,85
BMD_1966		conserved hypothetical protein	-1,66	-3,14	-3,04	-7,83
BMD_0368		intracellular protease, PfpI family	1,94	2,72	3,15	7,81
BMD_5178	speE	spermidine synthase	-1,54	-2,78	-3,49	-7,81
BMD_2245	proA	gamma-glutamyl phosphate reductase	2,29	3,29	2,23	7,80
BMD_2544		conserved hypothetical protein	1,33	3,98	2,48	7,79
BMD_1371	panB	3-methyl-2-oxobutanoate hydroxymethyltransferase	1,12	5,47	1,20	7,79
BMD_0728		conserved hypothetical protein	-1,84	-3,12	-2,81	-7,77
BMD_1694	dapA	dihydrodipicolinate synthase	1,01	5,34	1,42	7,77
BMD_5121	ssbB	single-strand binding protein	2,30	3,93	1,54	7,77
BMD_0702	appB	oligopeptide ABC transporter, permease protein AppB	1,91	3,62	2,24	7,77
BMD_4418	artP	arginine ABC transporter, arginine- binding protein ArtP	3,20	3,08	1,48	7,76
BMD_4211		transcription factor FapR	-1,37	-2,42	-3,90	-7,70
BMD_0536	pgcA	alpha-phosphoglucomutase	-1,92	-2,75	-3,02	-7,69
BMD_4459		proline dipeptidase	-1,62	-2,79	-3,28	-7,69
BMD_4966	yutF	4-nitrophenylphosphatase	-1,67	-2,45	-3,56	-7,69
BMD_3046		Glutathionylspermidine synthase	-1,67	-3,55	-2,46	-7,68
BMD_1509	fhuD	ferrichrome import ABC transporter, ferrichrome-binding protein FhuD	-2,11	-3,28	-2,29	-7,68
BMD_4185	hslU	heat shock protein HslVU, ATPase subunit HslU	-1,62	-2,52	-3,54	-7,68
BMD_1350	pycA	pyruvate carboxylase	-1,34	-1,96	-4,37	-7,67
BMD_0134	rplC	50S ribosomal protein L3	-1,51	-2,39	-3,77	-7,66
BMD_0572		hypothetical protein	1,61	3,04	3,00	7,65
BMD_1230	mtnA	methylthioribose-1-phosphate isomerase	2,87	3,35	1,42	7,64
BMD_0168		conserved hypothetical protein	1,47	2,39	3,78	7,63
BMD_4414		conserved hypothetical protein	1,78	2,88	2,97	7,63
BMD_4060		hypothetical protein	-1,91	-3,93	-1,78	-7,62
BMD_1963		GIY-YIG catalytic domain protein	-1,73	-2,86	-3,01	-7,60
BMD_2598	cbiG	cobalamin biosynthesis protein CbiG	-2,54	-3,95	-1,11	-7,60
BMD_2705		Alcohol dehydrogenase GroES-like domain/ zinc-binding dehydrogenase family oxidoreductase	-1,92	-4,01	-1,66	-7,59
BMD_3155	qoxB	cytochrome aa3 quinol oxidase, subunit I	-1,82	-3,07	-2,69	-7,58

Populationsheterogenität während der heterologen Proteinproduktion in *Bacillus megaterium*

BMD_0576		DNA-binding protein HU	1,63	3,79	2,14	7,56
BMD_1329	pdhD	pyruvate dehydrogenase complex E3 component, dihydrolipoamide dehydrogenase	-1,61	-2,48	-3,47	-7,56
BMD_1847	tal	transaldolase	-1,75	-2,81	-2,98	-7,54
BMD_4426	lpdV	dihydrolipoamide dehydrogenase E3 component of branched-chain alpha-keto acid dehydrogenase complex	-1,64	-2,84	-3,05	-7,53
BMD_2977		hypothetical protein	-1,62	-2,97	-2,92	-7,52
BMD_4424	bkdAB	2-oxoisovalerate dehydrogenase E1 component beta subunit	-1,72	-2,69	-3,11	-7,51
BMD_4242	pyrC	dihydroorotase	-1,65	-2,66	-3,17	-7,48
BMD_4978	sufS	cysteine desulfurase SufS	-1,78	-2,95	-2,73	-7,46
BMD_4959		NifU-like domain protein	-1,02	-1,88	-4,54	-7,45
BMD_4989		conserved hypothetical protein	-1,69	-3,16	-2,59	-7,44
BMD_2055	gltA	glutamate synthase, large subunit	-1,18	-1,39	-4,87	-7,44
BMD_2681		Oxidoreductase family protein	-2,32	-2,81	-2,30	-7,43
BMD_0838	gdh	Glucose 1-dehydrogenase	2,41	3,15	1,86	7,41
BMD_0013	guaB	inosine-5'-monophosphate dehydrogenase	1,06	-1,38	-4,96	-5,29
BMD_2227		conserved hypothetical protein	-1,79	-3,53	-2,05	-7,38
BMD_3767		small acid soluble spore protein	1,59	2,90	2,88	7,37
BMD_2374		conserved hypothetical protein	-1,69	-3,05	-2,62	-7,36
BMD_1604		cyanophycinase domain protein	1,53	2,02	3,81	7,36
BMD_0339	corA	magnesium and cobalt transport protein CorA	2,25	3,17	1,94	7,36
BMD_5139	atpE	ATP synthase F0, C subunit	-1,77	-2,48	-3,10	-7,35
BMD_1397	metC	cystathionine beta-lyase	2,25	3,67	1,39	7,32
BMD_1967		conserved hypothetical protein	-1,70	-3,05	-2,56	-7,31
BMD_1291		nitroreductase family protein	-1,60	-2,82	-2,89	-7,30
BMD_2439		conserved hypothetical protein	-1,54	-2,51	-3,25	-7,30
BMD_3156	qoxA	cytochrome aa3 quinol oxidase, subunit II	-1,62	-2,44	-3,20	-7,26
BMD_3245	cwlJ	cell wall hydrolase CwlJ	1,40	1,92	3,92	7,25
BMD_0788		ADP-ribose pyrophosphatase	1,49	2,04	3,70	7,24
BMD_2082		3-oxoacyl-[acyl-carrier-protein] synthase III protein 1	1,21	1,71	4,32	7,23
BMD_4232		protein of unknown function (DUF370)	-1,45	-2,63	-3,14	-7,23
BMD_0685		NAD dependent epimerase/dehydratase	2,23	3,36	1,64	7,23
BMD_5058	hisG	ATP phosphoribosyltransferase, catalytic subunit	-1,05	3,99	2,19	5,13
BMD_1284	ptsI	phosphotransferase system (PTS) enzyme I	-1,13	-1,38	-4,71	-7,22
BMD_2240		cobalamin binding domain protein	1,42	2,16	3,64	7,21
BMD_4314	ndk	nucleoside diphosphate kinase	-1,53	-1,85	-3,82	-7,21
BMD_4050		putative L-lysine 6-monooxygenase (NADPH)	-2,34	-2,73	-2,13	-7,20
BMD_0710	oppF	oligopeptide ABC transporter, ATP-binding protein OppF	-1,71	-1,13	-4,34	-7,18
BMD_3038		conserved hypothetical protein	-1,79	-3,00	-2,39	-7,18

BMD_1585	metN	methionine import ABC transporter, ATP-binding protein MetN	1,43	2,24	3,50	7,17
BMD_2963		putative metal ABC transporter, permease protein	-1,71	-3,15	-2,31	-7,17
BMD_5160	tal	transaldolase	-1,90	-2,69	-2,57	-7,16
BMD_4051		putative siderophore biosynthesis protein	-2,45	-2,94	-1,76	-7,15
BMD_5233		two-component response regulator	1,52	2,70	2,93	7,15
BMD_1921		hypothetical protein	1,11	1,97	4,06	7,14
BMD_0891	hxlA	3-hexulose-6-phosphate synthase	1,46	-1,79	-3,88	-4,21
BMD_1116		putative membrane protein	-1,38	-2,36	-3,38	-7,12
BMD_0294	gatB	glutamyl-tRNA(Gln) and/or aspartyl-tRNA(Asn)	-1,60	-2,70	-2,82	-7,11
BMD_3140		amidotransferase, B subunit				
BMD_1603	iadA	Lipase/Acylhydrolase (GDSL)	1,94	3,37	1,79	7,10
BMD_3992		beta-aspartyl peptidase	1,38	1,78	3,94	7,10
BMD_3149		putative membrane protein, YitT family	1,55	2,48	3,06	7,09
BMD_0069	ctc	aminoglycoside N(6')-acetyltransferase, GNAT family	1,65	3,14	2,29	7,08
BMD_4675	lonA	50S ribosomal protein L25/general stress protein Ctc	2,06	2,44	2,57	7,06
BMD_3852		ATP-dependent protease LonA	-1,57	-2,73	-2,76	-7,06
BMD_1560		sucrose-6-phosphate hydrolase (Sucrase) (Invertase)	-1,72	-2,83	-2,51	-7,06
BMD_4262	divIB	Amino acid permease family protein	1,68	3,83	1,55	7,06
BMD_1279	ribH	cell division initiation protein DivIB	-1,54	-2,58	-2,94	-7,05
BMD_5141	atpI	6,7-dimethyl-8-ribityllumazine synthase	-1,90	-2,79	-2,36	-7,05
BMD_0185	argI	ATP synthase F0, I subunit	-1,67	-2,17	-3,21	-7,05
BMD_4632		arginase	-2,44	-2,35	-2,25	-7,03
BMD_5134	atpD	conserved hypothetical protein	1,68	2,58	2,77	7,02
BMD_3709		ATP synthase F1, beta subunit	-2,24	-3,09	-1,69	-7,02
BMD_4458	efp	cell wall hydrolase cwlJ	1,35	1,81	3,85	7,01
BMD_1212	phaQ	translation elongation factor P	-1,35	-2,11	-3,55	-7,01
BMD_1125		poly-beta-hydroxybutyrate-responsive repressor	1,97	3,40	1,64	7,01
BMD_1310	dapH	putative tyrosine-protein phosphatase capC	1,86	3,57	1,57	7,00
BMD_0371		tetrahydrodipicolinate N-acetyltransferase	-1,52	-2,13	-3,34	-6,99
BMD_2468		conserved hypothetical protein	-1,41	-2,53	-3,04	-6,98
BMD_1705	nikD	conserved hypothetical protein ycsG	-1,44	-3,79	-1,73	-6,96
BMD_4151	tsf	nickel import ABC transporter, ATP-binding protein NikD	-2,10	-2,48	-2,38	-6,96
BMD_5091	gabP	translation elongation factor Ts	-1,24	-1,47	-4,24	-6,95
BMD_0707	oppB	GABA permease	-1,94	-2,92	-2,10	-6,95
		oligopeptide ABC transporter, permease protein OppB	-1,85	-1,37	-3,71	-6,93

BMD_4742	dnaB	replication initiation and membrane attachment protein (DnaB)	1,57	3,28	2,08	6,93
BMD_2087	pfkB	putative 1-phosphofructokinase	-1,67	-2,45	-2,81	-6,93
BMD_1591		conserved hypothetical protein	1,10	1,01	4,81	6,92
BMD_2871		conserved hypothetical protein	-1,77	-3,30	-1,85	-6,92
BMD_1013		CsbD-like protein	2,12	2,83	1,97	6,91
BMD_1211	phaP	polyhydroxyalkanoic acid inclusion protein PhaP	1,96	3,32	1,62	6,91
BMD_5177	speB	agmatinase	-1,47	-2,57	-2,86	-6,91
BMD_2088	fruA	phosphotransferase system (PTS) fructose-specific enzyme IIA/B component	-1,80	-2,73	-2,38	-6,90
BMD_1953	cobU	bifunctional adenosylcobalamin biosynthesis protein CobU	-1,73	-2,91	-2,25	-6,89
BMD_1365	dapB	dihydrodipicolinate reductase	-1,71	-2,96	-2,22	-6,89
BMD_4977	iscU	SUF system FeS assembly protein	-1,69	-2,38	-2,82	-6,89
BMD_1920		transcriptional regulator	-1,65	-2,44	-2,79	-6,88
BMD_3052	sipT	signal peptidase I T	-1,50	-2,76	-2,62	-6,88
BMD_0890		6-phospho 3-hexuloisomerase domain protein	1,49	-1,72	-3,68	-3,90
BMD_4288		putative methyltransferase	-1,76	-3,14	-1,96	-6,86
BMD_1954	cobS	adenosylcobinamide-GDP ribazoletransferase	-1,61	-2,97	-2,28	-6,86
BMD_1952	cobD	L-threonine-O-3-phosphate decarboxylase	-1,82	-2,70	-2,33	-6,85
BMD_4715	trx	thioredoxin	-1,24	-2,17	-3,43	-6,84
BMD_3160		hypothetical protein	1,59	2,10	3,14	6,83
BMD_4196	rplS	50S ribosomal protein L19	1,58	2,58	2,66	6,83
BMD_1442		conserved hypothetical protein	-1,61	-2,59	-2,62	-6,83
BMD_0126	rpoB	DNA-directed RNA polymerase, beta subunit	-1,22	-1,84	-3,76	-6,82
BMD_0719	mecA	competence-associated adapter protein	2,14	3,12	1,57	6,82
BMD_5104	comFA	ComFA competence protein	1,49	2,25	3,07	6,82
BMD_1347		conserved hypothetical protein	-1,76	-2,77	-2,29	-6,82
BMD_1460	dppE	dipeptide ABC transporter, dipeptide-binding protein DppE	1,83	2,73	2,25	6,81
BMD_0537	bioY	biotin biosynthesis protein BioY	1,74	2,91	2,16	6,81
BMD_3365		translocator protein, LysE family	-1,45	-2,90	-2,45	-6,80
BMD_1756	gerUC	spore germination protein	1,30	1,80	3,70	6,80
BMD_1412		conserved hypothetical protein	-1,48	-2,21	-3,10	-6,79
BMD_1237	mtnD	1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	2,40	3,25	-1,13	4,52
BMD_1571	levE	phosphotransferase system (PTS) fructose-specific enzyme IIB component	1,31	1,95	3,52	6,78
BMD_4835		putative lipoprotein	1,51	2,40	2,87	6,78
BMD_2226		ABC transporter, ATP-binding protein	-1,83	-3,11	-1,82	-6,77
BMD_0128	rplGB	50S ribosomal protein L7Ae	-1,31	-2,25	-3,21	-6,76
BMD_3676		transcriptional regulator, TetR family	1,83	3,01	1,91	6,76
BMD_3026		hypothetical protein	-1,28	-2,40	-3,09	-6,76

BMD_2438		conserved hypothetical protein	-1,64	-2,69	-2,43	-6,76
BMD_2922		putative pentachlorophenol-4-monooxygenase	1,38	1,64	3,74	6,75
BMD_1054		glucose 1-dehydrogenase	1,82	2,77	2,16	6,75
BMD_5057	hisD	histidinol dehydrogenase	-1,17	3,68	1,90	4,40
BMD_4249	ileS	isoleucyl-tRNA synthetase	-1,64	-2,29	-2,80	-6,73
BMD_1783		anti-sigma factor antagonist	1,85	2,98	1,89	6,72
BMD_3494		conserved hypothetical protein	1,63	2,15	2,94	6,72
BMD_3476		hypothetical protein	-1,49	-3,00	-2,22	-6,71
BMD_1706	nike	nickel import ABC transporter, ATP-binding protein Nike	-2,14	-2,59	-1,97	-6,71
BMD_4970	yutC	sporulation lipoprotein	1,28	1,77	3,66	6,71
BMD_2969		conserved hypothetical protein	-1,69	-2,82	-2,20	-6,71
BMD_2173	ykpA	ABC transporter, ATP-binding protein	-1,35	-2,07	-3,28	-6,70
BMD_1518		HTH-type transcriptional activator hxlR	1,34	2,01	3,33	6,68
BMD_3061		conserved hypothetical protein	1,46	2,85	2,36	6,67
BMD_0703	appC	oligopeptide ABC transporter, permease protein AppC	1,70	2,97	1,97	6,65
BMD_0722	pepF	oligoendopeptidase F	-1,68	-2,56	-2,41	-6,65
BMD_2527		acetyltransferase, GNAT family	-1,54	-2,53	-2,58	-6,65
BMD_4797	tyrS	tyrosyl-tRNA synthetase	-1,60	-2,64	-2,41	-6,64
BMD_4882	glgA	glycogen synthase	-1,92	-2,27	-2,45	-6,64
BMD_3374		glucose 1-dehydrogenase	1,61	2,69	2,33	6,64
BMD_2029		conserved hypothetical protein	-1,22	-2,29	-3,13	-6,64
BMD_0494	iolD	myo-inositol catabolism protein IolD	1,34	1,99	3,29	6,62
BMD_4233		conserved hypothetical protein	-1,36	-2,19	-3,07	-6,62
BMD_5059	hisZ	ATP phosphoribosyltransferase, regulatory subunit	-1,08	3,31	2,23	4,45
BMD_5156	tdk	thymidine kinase	-1,18	-1,82	-3,62	-6,62
BMD_1799		conserved hypothetical protein	2,21	2,94	1,48	6,62
BMD_4091	spoVK	stage V sporulation protein K	1,29	1,57	3,76	6,62
BMD_2938		putative branched-chain amino acid ABC transporter, ATP-binding protein	1,26	1,30	4,05	6,61
BMD_3262		hypothetical protein	-1,16	-2,34	-3,10	-6,60
BMD_4150	pyrH	uridylate kinase	-1,30	-1,77	-3,53	-6,60
BMD_0700	appF	oligopeptide ABC transporter, ATP-binding protein AppF	1,69	3,18	1,73	6,60
BMD_3677		multidrug resistance protein, SMR family	1,92	3,06	1,62	6,60
BMD_0132	tuf	translation elongation factor Tu (EF-Tu)	-1,25	-1,57	-3,77	-6,59
BMD_1846		hypothetical protein	-1,83	-2,90	-1,85	-6,58
BMD_1413		hypothetical protein	-1,25	-1,88	-3,44	-6,57
BMD_0201	mscL	large conductance mechanosensitive channel protein	-1,76	-2,83	-1,98	-6,57
BMD_1457	dppB	dipeptide ABC transporter, permease protein DppB	1,56	2,83	2,16	6,55
BMD_1825	brnQ	branched-chain amino acid transport system II carrier protein	1,25	2,43	2,87	6,55

BMD_1017		PTS system, diacetylchitobiose-specific IIB component	1,22	1,92	3,40	6,55
BMD_0708	oppC	oligopeptide ABC transporter, permease protein OppC	-1,80	-1,33	-3,42	-6,54
BMD_3089		conserved hypothetical protein	-1,48	-2,82	-2,25	-6,54
BMD_1059		hypothetical protein	1,72	2,60	2,22	6,54
BMD_4149	frr	ribosome recycling factor	-1,28	-1,72	-3,53	-6,53
BMD_0444		Serine protein kinase	1,34	1,68	3,49	6,51
BMD_4923	pgi	glucose-6-phosphate isomerase	-1,63	-2,67	-2,21	-6,50
BMD_4054		2,4-diaminobutyrate 4-transaminase	-2,25	-2,63	-1,61	-6,49
BMD_0729		putative GTP pyrophosphokinase	-1,54	-2,24	-2,70	-6,49
BMD_3811		amino acid permease	-1,61	-2,47	-2,40	-6,49
BMD_2930		NAD dependent epimerase/dehydratase family	1,45	2,24	2,79	6,48
BMD_3966		translocator protein, LysE family	-1,31	-2,89	-2,28	-6,48
BMD_2191		NADH dehydrogenase	-1,26	2,93	2,27	3,94
BMD_0169	cwlD	N-acetylmuramoyl-L-alanine amidase	1,42	2,03	3,02	6,47
BMD_2225		transcriptional regulator, gntR family protein	-1,59	-2,59	-2,28	-6,46
BMD_5052	hisI	histidine biosynthesis bifunctional protein HisI	-1,20	3,87	1,39	4,06
BMD_3234		AzlC family protein	-1,23	-2,41	-2,81	-6,45
BMD_3360		conserved hypothetical protein	1,62	1,85	2,98	6,44
BMD_4957	yutJ	NADH dehydrogenase YutJ	-1,67	-2,84	-1,93	-6,44
BMD_3130		hypothetical protein	-1,47	-2,61	-2,36	-6,44
BMD_0394		conserved hypothetical protein	1,75	2,71	1,98	6,44
BMD_4993		conserved hypothetical protein	1,41	2,16	2,87	6,44
BMD_0265		GMP synthase [glutamine-hydrolyzing]	-1,41	-2,17	-2,86	-6,44
BMD_4984	metN	methionine import ABC transporter, ATP-binding protein MetN	1,89	2,80	1,76	6,44
BMD_2682		transcriptional regulator, DeoR family	-2,21	-2,36	-1,86	-6,44
BMD_4477	comGE	prepilin-type N-terminal cleavage/methylation domain protein	1,61	2,32	2,51	6,43
BMD_1938	fhs	formate--tetrahydrofolate ligase	-1,63	-2,82	-1,98	-6,43
BMD_0664	rocB	arginine utilization protein RocB	1,65	2,35	2,43	6,43
BMD_4530		conserved hypothetical protein	-1,71	-2,28	-2,44	-6,43
BMD_3125		drug resistance MFS transporter, drug:H ⁺ antiporter-1 (14 Spanner) (DHA2) family	-1,66	-2,86	-1,91	-6,43
BMD_1119		glycosyl transferase, group 1	-1,57	-2,61	-2,24	-6,41
BMD_0370	ade	adenine deaminase	2,23	3,15	-1,03	4,34
BMD_5270	oxaA	membrane protein OxaA	-1,47	-2,20	-2,72	-6,40
BMD_1519	hxlA	3-hexulose-6-phosphate synthase	1,67	2,27	2,46	6,40
BMD_1997		hypothetical protein	1,81	2,64	1,94	6,39
BMD_1510	fhuC	ferrichrome import ABC transporter, ATP-binding protein FhuC	-1,83	-2,81	-1,75	-6,39

Populationsheterogenität während der heterologen Proteinproduktion in *Bacillus megaterium*

BMD_4027	argE	acetylornithine deacetylase	1,28	1,79	3,32	6,39
BMD_5006	ycel	major facilitator family transporter	1,23	2,83	2,32	6,38
BMD_2761		hypothetical protein	-2,02	-2,51	-1,84	-6,38
BMD_5166		methylmalonyl-CoA mutase	1,32	2,22	2,81	6,36
BMD_4494		nucleotidyl transferase family	1,34	1,93	3,09	6,36
BMD_2617		conserved hypothetical protein	-1,82	-2,59	-1,95	-6,36
BMD_0122	rplA	50S ribosomal protein L1	-1,06	-1,34	-3,95	-6,35
BMD_1392		hypothetical protein	1,34	2,72	2,29	6,35
BMD_3428		hypothetical protein	-1,55	-2,74	-2,05	-6,35
BMD_2460	cwlO	secreted cell wall DL- endopeptidase	-1,66	-3,30	-1,38	-6,33
BMD_1277	ribE	riboflavin synthase, alpha subunit	-1,70	-2,47	-2,16	-6,33
BMD_0125		16S rRNA m(2)G 1207 methyltransferase	-1,34	-2,19	-2,78	-6,31
BMD_3407		Proton/ glutamate symport protein	1,19	1,40	3,71	6,30
BMD_0230	rsbX	phosphoserine phosphatase RsbX	1,92	3,02	1,36	6,30
BMD_3115		malate dehydrogenase	-1,51	-1,90	-2,88	-6,30
BMD_4564	comER	ComE operon protein	1,49	2,39	2,40	6,28
BMD_1769		hut operon positive regulatory protein	1,59	2,28	2,39	6,27
BMD_1215	phaB	acetoacetyl-CoA reductase	-1,54	-1,16	-3,55	-6,26
BMD_0098	lysS	lysyl-tRNA synthetase	-1,35	-2,03	-2,87	-6,26
BMD_3755	sspA	small acid soluble spore protein	1,54	2,67	2,04	6,26
BMD_0110	ispF	2C-methyl-D-erythritol 2,4- cyclodiphosphate synthase	-1,59	-2,55	-2,11	-6,25
BMD_4231	gmk	guanylate kinase	-1,21	-1,77	-3,27	-6,25
BMD_3141		transporter, AcrB/AcrD/AcrF family	1,92	3,28	1,04	6,24
BMD_1201	msrAB	peptide methionine sulfoxide reductase MsrA/MsrB	-1,41	-2,10	-2,73	-6,23
BMD_1388		hypothetical protein	1,83	2,04	2,36	6,23
BMD_1126	galU	UTP-glucose-1-phosphate uridylyltransferase	1,74	3,06	1,43	6,23
BMD_5143	vpr	minor extracellular protease Vpr	1,43	2,04	2,75	6,22
BMD_4167	cheY	chemotaxis protein CheY	1,36	2,07	2,80	6,22
BMD_0570		conserved hypothetical protein	-1,28	-1,85	-3,08	-6,21
BMD_3806		hypothetical protein	1,46	1,94	2,80	6,20
BMD_1662		transcriptional regulator, LysR family	-1,33	-2,36	-2,51	-6,20
BMD_3908		hypothetical protein	1,84	3,07	1,28	6,20
BMD_2170		conserved hypothetical protein	-1,54	-2,61	-2,04	-6,20
BMD_0561	sspC	small acid soluble spore protein C	1,33	2,15	2,72	6,19
BMD_3577		2-amino-3-carboxymuconate-6- semialdehyde decarboxylase	1,24	1,61	3,34	6,19
BMD_5142		putative membrane protein	-1,67	-2,29	-2,22	-6,18
BMD_2379		rhodanese domain protein	-1,36	-2,47	-2,34	-6,17
BMD_0328		putative nickel transporter	-1,83	-2,37	-1,97	-6,16
BMD_1788		conserved hypothetical protein	1,55	2,27	2,34	6,16
BMD_4146	dxr	1-deoxy-D-xylulose 5-phosphate reductoisomerase	-1,41	-2,10	-2,64	-6,16
BMD_5168	acdA	acyl-CoA dehydrogenase	1,26	1,93	2,96	6,16
BMD_1782	ysnF	stress response protein YsnF	1,80	2,68	1,67	6,16
BMD_0261	groEL	60 kDa chaperonin	-1,28	-1,22	-3,65	-6,15

Populationsheterogenität während der heterologen Proteinproduktion in *Bacillus megaterium*

BMD_5258	engD	GTP-binding protein EngD	-1,26	-2,09	-2,80	-6,15
BMD_1117		glycosyl transferase, family 2	-1,47	-2,37	-2,31	-6,15
BMD_5205		endonuclease/exonuclease/phosphatase family protein	1,94	2,92	1,28	6,15
BMD_4696		protein of unknown function (DUF47)	-1,39	-2,14	-2,61	-6,14
BMD_2723		conserved hypothetical protein	-1,37	-2,53	-2,23	-6,14
BMD_0335		Sodium:alanine symporter family	1,79	2,50	1,84	6,14
BMD_4954		iron-sulfur cluster assembly accessory protein	-1,14	-1,79	-3,19	-6,13
BMD_5183		conserved hypothetical protein	1,18	1,35	3,59	6,13
BMD_4240	pyrAB	carbamoyl-phosphate synthase, large subunit	-1,68	-2,32	-2,13	-6,12
BMD_2025		Riboflavin biosynthesis protein, Riboflavin kinase	1,16	1,51	3,45	6,12
BMD_5196	zwf	glucose-6-phosphate 1-dehydrogenase	-1,34	-2,04	-2,74	-6,12
BMD_1313		mechanosensitive ion channel	2,05	2,68	1,38	6,11
BMD_1200		transcriptional regulator, MarR family	-1,44	-2,40	-2,26	-6,10
BMD_2784		DNA-binding protein HU	-1,64	-2,75	-1,71	-6,10
BMD_1542		Hut operon positive regulatory protein	-1,60	-2,39	-2,11	-6,10
BMD_4087	glnR	glutamine synthetase repressor	-1,28	-2,17	-2,64	-6,09
BMD_4272		conserved hypothetical protein	-1,43	-2,04	-2,62	-6,09
BMD_2463	accC	acetyl-CoA carboxylase, biotin carboxylase	-1,51	-3,34	-1,24	-6,09
BMD_0325		putative ABC transporter, ATP-binding protein	1,35	1,96	2,77	6,09
BMD_4290		conserved hypothetical protein	-1,63	-2,85	-1,61	-6,08
BMD_5111		membrane bound transcriptional regulator	1,45	2,21	2,42	6,08
BMD_2117		general stress protein	1,66	1,84	2,58	6,08
BMD_4976	sufB	FeS assembly protein SufB	-1,53	-2,08	-2,47	-6,08
BMD_1244		hypothetical protein	-1,47	-2,45	-2,16	-6,08
BMD_2949		hydrolase, alpha/beta fold family	-1,53	-2,48	-2,06	-6,07
BMD_2021	tcyK	sulfur-containing amino acid ABC transporter, sulfur-containing amino acid-binding protein TcyK	1,12	1,49	3,46	6,07
BMD_3625		two-component response regulator	1,23	1,61	3,23	6,07
BMD_0895		conserved hypothetical protein	-1,87	-2,08	-2,11	-6,06
BMD_3379		LemA family protein	-1,46	-2,41	-2,18	-6,05
BMD_3200		phage-related replication protein	1,34	1,90	2,81	6,05
BMD_4596		protein of unknown function (DUF965)	-1,30	-1,76	-2,98	-6,05
BMD_0197		PRD domain	1,31	2,00	2,73	6,05
BMD_4807		conserved hypothetical protein	1,65	2,33	2,07	6,05
BMD_2991	ureA	urease, gamma subunit	-1,72	-2,76	-1,56	-6,04
BMD_1367		conserved hypothetical protein	-1,55	-2,45	-2,03	-6,04
BMD_2615	amt	ammonium transporter	-1,72	-2,67	-1,64	-6,03
BMD_0595	ecsC	ABC transporter, substrate-binding protein EcsC	-1,42	-1,71	-2,90	-6,03

BMD_3621		Oxidoreductase family, NAD-binding Rossmann fold	1,27	1,93	2,83	6,03
BMD_4481	comGA	ComG operon protein	1,29	2,04	2,70	6,03
BMD_0017	serS	seryl-tRNA synthetase	-1,47	-2,27	-2,29	-6,03
BMD_3812		GABA permease	-1,70	-2,55	-1,77	-6,03
BMD_0159	rpsK	30S ribosomal protein S11	-1,35	-1,96	-2,71	-6,03
BMD_2939		putative branched-chain amino acid ABC transporter, ATP-binding protein	1,17	1,48	3,37	6,02
BMD_4442	ispA	geranyltranstransferase	-1,43	-2,22	-2,37	-6,02
BMD_2465		Allophanate hydrolase subunit 2	-1,47	-2,88	-1,66	-6,02
BMD_4411		peptidase T	-1,60	-2,02	-2,40	-6,02
BMD_1366	mgsA	methylglyoxal synthase	-1,60	-2,42	-2,00	-6,02
BMD_3813	putC	putative 1-pyrroline-5-carboxylate dehydrogenase	-1,66	-2,35	-2,01	-6,02
BMD_2469		conserved hypothetical protein	-1,42	-2,99	-1,59	-6,01
BMD_1358		conserved hypothetical protein	1,54	2,66	1,80	6,01
BMD_4678	tig	trigger factor	-1,05	-1,37	-3,58	-6,00
BMD_1948		Thioredoxin	-1,09	-1,60	-3,31	-6,00
BMD_4427	buk	butyrate kinase	-1,46	-2,25	-2,29	-6,00
BMD_1461		putative L,D-carboxypeptidase	1,46	1,65	2,89	6,00
BMD_3798	clpP	ATP-dependent Clp protease, proteolytic subunit ClpP	1,33	1,37	3,30	6,00
BMD_0781	speE	spermidine synthase	1,27	1,70	3,02	6,00
BMD_3041		hypothetical protein	1,77	2,33	1,90	6,00
BMD_1860	xylT	xylose permease	1,43	1,76	1,18	4,37

Danksagung

Es war nicht immer einfach, die Dissertation und die Familie unter einen Hut zu bekommen – ich habe derzeit wirklich keine Ahnung, wo ich die ganze Steine hin packen soll, die ich im Laufe dieser Zeit versucht habe, aus dem Weg zu räumen.

Daher gilt mein Dank all jenen, die dies erkannt haben, aufmunternde Worte fanden und mit denen man sich organisatorisch abwechseln konnte.

Des Weiteren möchte ich den Personen meinen Dank aussprechen,

.... die meine Arbeit gewürdigt haben.

... die mich vermisst haben, wenn ich früher von einer Party gegangen bin, um im Labor noch was zu erledigen.

... die verwundert waren, wie früh ein Arbeitstag beginnen kann – eine Uhrzeit, bei der alle Hühner und Hähne noch schlafen.

Zu guter Letzt möchte ich mich bei den Menschen bedanken, die die arbeitsreiche Zeit gesehen haben und erkannt haben, dass es das Projekt dahin gebracht hat, zu dem es wurde.